

21ST international conference on science
and technology indicators



STI Conference 2016 · València

Peripheries, frontiers and beyond

14 · 16 September 2016
Universitat Politècnica de València

BOOK OF PROCEEDINGS

www.sti2016.org

Congress UPV

Proceedings of the 21ST International Conference on Science and Technology Indicators

The contents of this publication have been evaluated by the Scientific Committee which it relates and the procedure set out <http://ocs.editorial.upv.es/index.php/STI2016/STI2016>

Scientific Editors

Ismael Ràfols
Jordi Molas-Gallart
Elena Castro-Martínez
Richard Woolley

© of the text: the authors

© 2016, Editorial Universitat Politècnica de València
www.lalibreria.upv.es / Ref.: 6333_01_01_01

ISBN: 978-84-9048-519-4 (print version)
Print on-demand

DOI: <http://dx.doi.org/10.4995/STI2016.2016.4543>



Proceedings of the 21ST International Conference on Science and Technology Indicators
This book is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs-4.0 International License
Editorial Universitat Politècnica de València

TABLE OF CONTENTS

	Page
INTRODUCTION.....	9
CONFERENCE THEME.....	10
COMMITTEES.....	11
SPONSORS.....	13
CONFERENCE PROGRAMME.....	14
PLENARIES.....	40
ORAL PRESENTATIONS	
CHAPTER 1: Data Infrastructure and Data Quality	
• Must Metrics Serve the Audit Society? Addressing Marketization in Open Access Publishing and Humanities Analytics – <i>Newfield, Christopher; Muellerleile, Christopher</i> † (session 2.1)	51
• On the Peripheries of Scholarly Infrastructure: A look at the Journals Using Open Journal Systems – <i>Alperin, Juan Pablo; Stranack, Kevin; Garnett, Alex</i> † (session 2.4)	58
• Why researchers publish in journals not indexed in mainstream databases: training, bridging and gap-filling – <i>Chavarro, Diego; Tang, Puay; Rafols, Ismael</i> † (session 2.4)	66
• Identifying Sources of Scientific Knowledge: classifying non-source items in the WoS – <i>Calero-Medina, Clara M.</i> † (session 2.4)	73
• Data quality and consistency in Scopus and Web of Science in their indexing of Czech Journals – <i>Mika, Pavel; Szarzec, Jakub; Sivertsen, Gunnar</i> † (session 3.4)	78
• Missing citations due to exact reference matching: Analysis of a random sample from WoS. Are publications from peripheral countries disadvantaged? – <i>Donner, Paul</i> † (session 3.4)	85
• Funding Acknowledgements in the Web of Science: inconsistencies in data collection and standardization of funding organizations – <i>van Honk, Jeroen; Calero-Medina, Clara; Costas, Rodrigo</i> † (session 3.4)	90
• Open Data in Global Environmental Research: Findings from the Community – <i>Schmidt, Birgit; Gemeinholzer, Birgit; Treolar, Andrew</i> † (session 3.4)	97
• Examining data access and use in science – <i>Yan, Erjia; Zhao, Mengnan</i> † (session 4.4)	100
• SMS: a linked open data infrastructure for science and innovation studies – <i>Van den Besselaar, Peter; Khalili, Ali; Idrissou, Al; Loizou, Antonis; Schlobach, Stefan; van Harmelen, Frank</i> † (session 4.4)	106
• Data Citation Policies of Data Providers within the scope of Longitudinal Studies in Life Course Research – <i>Reinhold, Anke; Rittberger, Marc; Mahrholz, Nadine</i> † (session 4.4)	115
• Stepping up Information Infrastructures and Statistical Reporting – Monitoring the German Excellence Initiative – <i>Reinhardt, Anke</i> † (session 4.4)	122
CHAPTER 2: Smart Use of Indicators for Innovation Policy	
• Innovation indicators: Towards a User’s guide – <i>Iizuka, Michiko; Hollanders, Hugo</i> † (session 1.1)	127
• Analysing innovation policy indicators through a functional approach: the aeronautic industry case – <i>Haddad, Carolina R.; Uriona Maldonado, Mauricio</i> † (session 1.1)	139
• Assessing the performance of national innovation systems in Europe – <i>Edquist, Charles; Zabala-Iturriagagoitia, Jon Mikel</i> † (session 1.1)	149
• The process of construction of evidences: An analysis of the use of indicators in two decisions of innovation policy – <i>Boavida, Nuno; Moniz, António; Laranja, Manuel</i> † (session 1.2)	156
• Evidence-based policy learning: the case of the Research Excellence Indicator – <i>Hardeman, Sjoerd; Vertesy, Daniel</i> † (session 2.3)	168
• Who sets up the bridge? Tracking scientific collaborations between China and the EU28 – <i>Wang, Lili; Wang, Xianwen</i> † (session 2.3)	178
• A case study about the Colombian Observatory of Science and Technology: between context relevant and internationally comparable indicators – <i>Salazar, Monica; Crespi, Gustavo</i> † (session 2.3)	187
• Lessons learned for priority setting and indicators relevant to the impact of research programmes in Europe and Emerging Economies. An evidence-based debate between the research and the policy-shaping community – <i>Caloghirou, Yannis; Vonortas, Nicolas</i> † (session 4.1)	192

• Thirty years of European Collaboration in R&D: a Search for Key Network Indicators – <i>Caloghirou, Yannis; Protogerou, Aimilia; Siokas, Evangelos</i> † (session 4.1)	197
• Performance indicators for areas of innovation: international perspective – <i>Plonski, Guilherme Ary; Getz, Daphne; Barge-Gil, Andres; Modrego Rico, Aurelia</i> † (session 5.4)	208

CHAPTER 3: Measuring Innovation

• Measuring macro-level effects of the global economic recession on university-industry research cooperation – <i>Azagra-Caro, Joaquín M.; Tijssen, Robert J.W.; Yegros-Yegros, Alfredo</i> † (session 3.2)	213
• The impact of methodology in innovation measurement – <i>Wilhelmsen, Lars; Bugge, Markus; Solberg, Espen</i> † (session 3.5)	222
• A Critical Assessment of the Quality and Validity of Composite Indicators of Innovation – <i>Vértesy, Dániel</i> † (session 3.5)	229
• Innovation strategies in Latin American firms – <i>Vargas, Fernando</i> † (session 4.5)	234
• Innovation dynamics of Salvadoran agri-food industry from an evolutionary perspective – <i>Peraza Castaneda, Elias Humberto; Aleixandre Mendizábal, Guillermo</i> † (session 4.5)	246
• Elucidate Innovation Performance of Technology-driven Mergers and Acquisitions – <i>Huang, Lu; Wang, Kangrui; Yu, Huizhu; Shang, Lining; Mitkova, Liliana</i> † (session 4.5)	252
• Measuring Innovation in Government – <i>Mahroum, Sami</i> † (session 7.1)	269
• Measuring Innovation in Government: An International Review and Case Study of the UAE – <i>Mahroum, Sami; Al-Saleh, Yasser; Al-Hashmi, Shatha</i> † (session 4.5)	272
• Baseline of indicators for R&D and Innovation in ICT: a tool for decision-making, design and monitoring of public policies – <i>Mora Holguín, Henry; Lucio-Arias, Diana; Zárate, Sandra; Castro, Nayibe; Pardo, Clara</i> † (session 6.7)	283
• Measuring originality: common patterns of invention in research and technology organizations – <i>Tang, David Li; Wiseman, Erica; Keating, Tamara; Archambeault, Jean</i> † (session 6.7)	289
• Linking international trademark databases to inform IP research and policy – <i>Petrie, Stephen</i> † (session 6.7)	298
• Detecting emerging trends and country specializations in Energy Efficiency – <i>De Filippo, Daniela; Pandiella-Dominique, Andres</i> † (session 6.7)	303

CHAPTER 4: Societal Impact

• Unveiling Research Agendas: a study of the influences on research problem selection among academic researchers – <i>Bianco, Mariela; Sutz, Judith</i> † (session 2.1)	313
• “If we come out with the wrong answer that may affect investments”: Exploring how evaluators were influenced by political considerations during the assessment of societal impact – <i>Samuel, GS.; Derrick, GE.</i> † (session 2.1)	318
• ‘Productive interactions’ for societal impact: Developing a research information system for agriculture (RIS-Agric) at Stellenbosch University, South Africa – <i>Boshoff, Nelius; Esterhuysen, Harrie</i> † (session 3.1)	325
• Publication patterns in UK research assessment 1992-2013 – <i>Adams, Jonathan; Loach, Tamar; Szomszor, Martin</i> † (session 3.1)	332
• Societal impact metrics for non-patentable research in dentistry – <i>Hicks, Diana; Isett, Kim; Melkers, Julia; Song, Le; Trivedi, Rakshit</i> † (session 4.2)	343
• The Evolution of Scientific Trajectories in Rice: Mapping the Relation between Research and Societal Priorities – <i>Ciarli, Tommaso; Rafols, Ismael</i> † (session 4.2)	347
• Research Quality Plus (RQ+). A Holistic Approach to Evaluating Research – <i>McLean, Robert K.D.; Feinstein, Osvaldo N.</i> † (session 4.2)	357
• Impact of Research on Development in Cameroon: convergence between supply and research needs in the food sector – <i>Minkoua, Jules René; Temple, Ludovic</i> † (session 5.2)	366
• Monitoring the Evolution and Benefits of Responsible Research and Innovation (MoRRI) – a preliminary framework for measuring RRI dimensions – <i>Mejlgaard, Niels; Buehrer, Susanne; Griessler, Erich; Lindner, Ralf; Maroulis, Nikos; Meijer, Ingeborg; Peter, Viola; Rafols, Ismael; Ravn, Tine; Stilgoe, Jack; Tsipouri, Lena; Woolley, Richard; Wroblewski, Angela</i> † (session 5.2)	375
• “All this grassroots, real life knowledge”: Comparing perceived with realised concerns of including non-academic evaluators in societal impact assessment – <i>Derrick, Gemma; Samuel, GS.</i> † (session 5.2)	379

CHAPTER 5: Mission Oriented Research

• Using novel computer-assisted linguistic analysis techniques to assess the timeliness and impact of FP7 Health's research – a work in progress report – <i>Stanciauskas, Vilijus; Brozaitis, Haroldas; Manola, Natalia; Metaxas, Omiros; Galsworthy, Mike</i> † (session 5.5)	388
• Professional impact of clinical research – <i>Nelhans, Gustaf</i> † (session 5.5)	393

• Technology Push / Market Pull Indicators in Healthcare – <i>Efimenko, Irina; Khoroshevsky, Vladimir; Noyons, Ed; Nochevkin, Evgeny</i> † (session 5.5)	401
• Mapping the networks of cancer research in Portugal: first results – <i>Brás, Oriana Rainho; Cointet, Jean-Philippe; Nunes, João A.; David, Leonor; Cambrosio, Alberto</i> † (session 5.5)	406
• Access to global health research. Prevalence and cost of gold and hybrid open access – <i>Haustein, Stefanie; Smith, Elise; Mongeon, Philippe; Shu, Fei; Larivière, Vincent</i> † (session 6.5)	410
• Scientific research on diseases: the distinct profile of developed and developing countries – <i>Yegros-Yegros, Alfredo; Rafols, Ismael; Abad-García, María F.; Mugnaini, Rogério; Meijer, Ingeborg</i> † (session 6.5)	419
• Biodiversity Sustainability of Phytomedicine Research: a 3-dimensions analysis around the North-South divide – <i>Gorry, Philippe</i> † (session 6.5)	424
• In Re The Academic Cartography of Sugar Sweetened Beverages: Scientific and Technical Information, Interdisciplinarity, and Legal Academia – <i>White, Lexi C.</i> † (session 6.5)	435
• Metrics and Agricultural Science - measuring Multidisciplinary and Applied Research – <i>Holt, Ilkay</i> † (session 7.2)	439
• Network analysis to support research management: evidence from the Fiocruz Observatory in Science, Technology and Innovation in Health – <i>Fonseca, Bruna; Sampaio, Ricardo B.; Silva, Marcus V.; Dos Santos, Paula X.</i> † (session 8.5)	442
• Partial alphabetical authorship in medical research: an exploratory analysis – <i>Mongeon, Philippe; Smith, Elise; Joyal, Bruno; Larivière, Vincent</i> † (session 8.5)	448
• The bibliometric behaviour of an expanding specialisation of medical research – <i>Thelwall, Mike; Levitt, Jonathan</i> † (session 8.5)	453

CHAPTER 6: Rethinking Evaluation

• Outlining an analytical framework for mapping research evaluation landscapes – <i>Åström, Fredrik</i> † (session 1.2)	462
• When the Brightest are not the Best – <i>Valente, Marco</i> † (session 1.2)	467
• The reward (eco)system of science: More than the sum of its parts? – <i>Desrochers, Nadine; Haustein, Stefanie; Alperin, Juan Pablo; Bowman, Timothy D.; Díaz-Faes, Adrián A.; Larivière, Vincent; Mongeon, Philippe; Paul-Hus, Adèle; Quan-Haase, Anabel; Smith, Elise; Sugimoto, Cassidy R.</i> † (session 1.3)	477
• Capturing knowledge integration through collaborations: measures of the diversity and coherence in multiple proximity dimensions – <i>Lang, Frédérique; Rafols, Ismael; Hopkins, Michael</i> † (session 1.4)	485
• Using a network-based approach to identify interactions structure for innovation in a low-technology intensive sector – <i>Aouinaït, Camille; Lepori, Benedetto; Christen, Danilo; Carlen, Christoph; Foray, Dominique</i> † (session 1.4)	497
• Interdisciplinary and transdisciplinary institutions: do they constitute peripheries among cultures? – <i>Vienni, Bianca; Vilsmaier, Ulli</i> † (session 1.4)	506
• "Putting in more than you take out" Towards evaluating research based on its public (not private) contributions – <i>Benneworth, Paul; Olmos-Peñuela, Julia; Castro-Martínez, Elena</i> † (session 1.4)	511
• On multiplying methods in the field of research evaluation – <i>Derrick, Gemma; Molas-Gallart, Jordi; de Rijcke, Sarah; Meijer, Ingeborg; van der Weijden, Inge; Wouters, Paul</i> † (session 2.2)	516
• Why DORA does not stand a chance in the biosciences – <i>Gläser, Jochen</i> † (session 8.1)	518
• Are institutional missions aligned with journal-based or document-based disciplinary structures? – <i>Klavans, Richard; Boyack, Kevin W.</i> † (session 8.1)	523
• Science policy through stimulating scholarly output. Reanalyzing the Australian case – <i>Van den Besselaar, Peter; Heyman, Ulf; Sandström, Ulf</i> † (session 8.1)	531
• The need for contextualized scientometric analysis: An opinion paper – <i>Waltman, Ludo; van Eck, Nees Jan</i> † (session 8.1)	541
• Fake Academic Degrees as an Indicator for Severe Reputation Crisis in the Scientific Community of Russian Federation – <i>Rostovtsev, Andrey; Kostinskiy, Alexander</i> † (session 8.2)	550

CHAPTER 7: Programme and Project Assessment

• An assessment of EU-funded research projects: innovators and their innovative potential – <i>Nepelski, Daniel; Pesole, Annarosa; Van Roy, Vincent; O'Neill, Eoghan</i> † (session 6.6)	555
• Evaluating the impact of public space investments with limited time and funds: (methodological) lessons from a Swiss case study – <i>Barjak, Franz</i> † (session 6.6)	562
• Researchers and Institutions in the Periphery: Challenges in Measuring Research Capacity for Geographically Specific Programs in the U.S. – <i>Melkers, Julia</i> † (session 6.6)	571
• Assessing marine biotechnology research centres in peripheral regions: developing global and local STI indicators – <i>Robinson, Douglas K.R.; Schoen, Antoine; Laurens, Patricia; Horellou, Sülzel; Colas, Pierre; Larédo, Philippe</i> † (session 6.6)	575
• Examining to what extent the source of competitive funding influences scientific impact. A case study of Danish EU FP7 funded projects – <i>Ryan, Thomas K.; Schneider, Jesper W.</i> † (session 4.8)	581

• The Determinants of National Funding in Trans-national Joint Research: Exploring the Proximity Dimensions – Reale, Emanuela; Spinello, Andrea; Zinilli, Antonio (session 4.8)	590
• Beyond funding: What can acknowledgements reveal about credit distribution in science? – Paul-Hus, Adèle; Díaz-Faes, Adrián; Desrochers, Nadine; Costas, Rodrigo; Sainte-Marie, Maxime; Macaluso, Benoît; Larivière, Vincent (session 4.8)	598
• Allocating organisational level funding on the basis of Research Performance Based assessments, a comparative analysis of the EU Member States in international perspective – Jonkers, Koen; Zacharewicz, Thomas; Lepori, Benedetto; Reale, Emanuela (session 4.8)	607

CHAPTER 8: Collaborations, Mobility and Internationalisation

• Predicting STEM Career Success by STI Knowledge Utilization Patterns – Bozeman, Barry; Youtie, Jan; Bretschneider, Stuart (session 5.3)	617
• The Credibility of Policy Reporting Across Learning Disciplines – Carley, Stephen; Youtie, Jan; Solomon, Gregg; Porter, Alan (session 5.3)	622
• Mobility in the academic careers at the Flemish universities – Results from the Human Resources in Research database – Debacker, Noëmi; Vandeveldde, Karen (session 6.4)	627
• Gatekeeping African studies: A preliminary insight on what do editorial boards indicate about the nature and structure of research brokerage – Mendonça, Sandro; Pereira, João; Ferreira, Manuel E. (session 6.4)	635
• The Network of International Student Mobility – Voegtle, Eva; Windzio, Michael (session 7.4)	637
• Scientific mobility of Early Career Researchers in Spain and The Netherlands through their publications – Robinson-Garcia, Nicolas; Cañibano, Carolina; Woolley, Richard; Costas, Rodrigo (session 7.4)	644
• Big Science, co-publication and collaboration: getting to the core – Kahn, Michael (session 7.4)	653
• Autonomy vs. dependency of scientific collaboration in scientific performance – Chinchilla-Rodríguez, Zaida; Miguel, Sandra; Perianes-Rodríguez, Antonio; O valle-Perandones, María-Antonia; Olmeda-Gómez, Carlos (session 7.4)	662
• The world network of scientific collaborations between cities: domestic or international dynamics? – Maisonobe, Marion; Eckert, Denis; Grossetti, Michel; Jégou, Laurent; Milard, Béatrice (session 8.4)	671
• Trends in the inter-regional and international research collaboration of the PRC's regions: 2000-2015 – Luwel, Marc; van Wijk, Erik; Van der Wurff, Lambertus J.; Wang, Lili (session 8.4)	682
• Iran's scientific dominance and the emergence of South-East Asian countries in the Arab Gulf Region – Moed, Henk (session 8.4)	693
• How international is internationally collaborated research? A bibliometric study of Russian surname holder collaboration networks – Karaulova, Maria; Gök, Abdullah; Shapira, Philip (session 8.4)	702

CHAPTER 9: Careers, Labour Market and Individual Performance

• Information sources – information targets: evaluative aspects of the scientists' publication strategies – Glänzel, Wolfgang; Chi, Pei-Shan; Gumpenberger, Christian; Gorraiz, Juan (session 4.7)	708
• The Effect of Holding a Research Chair on Scientists' Impact – Mirnezami, Seyed Reza; Beaudry, Catherine (session 4.7)	718
• Public-private collaboration and scientific impact: an analysis at the level of the individual researcher – Bloch, Carter; Ryan, Thomas K.; Andersen, Jens Peter (session 4.7)	733
• Survey on the Labour Market Position of PhD Graduates: Competence comparison and relation between PhD and current employment – Heuritsch, Julia; Waaijer, Cathelijn J. F.; van der Weijden, Inge C. M. (session 5.1)	741
• Beyond the indicators: formulation of the career strategies of scientists – Pálinkó, Éva (session 5.1)	750
• Exploring predictors of scientific performance with decision tree analysis: The case of research excellence in early career mathematics – Lindahl, Jonas (session 5.1)	759
• Stability and longevity in the publication careers of U.S. doctorate recipients – Waaijer, Cathelijn J. F.; Macaluso, Benoît; Sugimoto, Cassidy R.; Larivière, Vincent (session 5.1)	766
• Developing research career indicators using open data: the RISIS infrastructure – Cañibano, Carolina; Woolley, Richard; Iversen, Eric; Hinze, Sybille; Hornbostel, Stefan; Tesch, Jakob (session 5.2)	771
• On the extent of researcher mobility and indicators for mobility – Barriere, Stina; Baard, Patrik; Karlsson, Staffan; Jacobsson, Carl; Tumpene, John (session 5.2)	780
• Progress on mobility and instability of research personnel in Japan: scientometrics on a job-posting database for monitoring the academic job market – Kawashima, Hirotaka; Yamashita, Yasuhiro (session 5.2)	789
• National and international scientific elites: an analysis of Chinese scholars – Shu, Fei; Larivière, Vincent; Julien, Charles-Antoine (session 5.2)	797

CHAPTER 10: Gender

• Gender in science: a periphery? – <i>van der Weijden, Inge; Reale, Emanuela; Meijer, Ingeborg</i> (session 1.6)	803
• What drives the gender gap in STEM? The SAGA Science, Technology and Innovation Gender Objectives List (STI GOL) as a new approach to linking indicators to STI policies – <i>Fernández, Ernesto; Schaaper, Martin; Bello, Alessandro</i> (session 2.6)	808
• Picking the best publications to showcase graduate courses: Do institutional mechanisms reinforce gender differences? – <i>Leta, Jacqueline; Cabanac, Guillaume</i> (session 2.6)	812
• What factors influence scientific and technological output: The case of Thailand and Malaysia – <i>Beaudry, Catherine; St-Pierre, Carl</i> (session 2.6)	819
• Identifying the Gender Dimension in Research Content – <i>Campbell, David; Lalonde, Bastien St. Louis; Tippett, Chantale; Archambault, Éric; Callaert, Julie; Mantouvalou, Katerina; Arora, Lucy</i> (session 3.6)	832
• Gender Differences in Synchronous and Diachronous Self-citations – <i>Ghiasi, Gita; Larivière, Vincent; Sugimoto, Cassidy</i> (session 3.6)	844
• Mapping the author gender-distribution of disease-specific medical research – <i>Andersen, Jens Peter; Schneider, Jesper Wiborg; Nielsen, Mathias Wullum</i> (session 3.6)	852
• Indicators for constructing scientific excellence: 'Independence' in the ERC Starting Grant – <i>Schiffbaenker, Helene; Holzinger, Florian</i> (session 3.6)	863

CHAPTER 11: Social Sciences and Humanities

• Indicators for Research Performance in the Humanities? The Scholars' View on Research Quality and Indicators – <i>Ochsner, Michael; Hug, Sven E.</i> (session 3.3)	873
• Quality criteria and indicators for research in Theology – What to do with quantitative measures? – <i>Martens, Silvia; Schatz, Wolfgang</i> (session 3.3)	881
• Social Impact Open Repository (SIOR). Transforming the peripheral space of social impact of research – <i>Joanpere, Mar; Samano, Elvira; Gómez, Aitor</i> (session 4.3)	884
• Je veux bien, mais me citez-vous? On publication language strategies in an anglicized research landscape – <i>Desrochers, Nadine; Larivière, Vincent</i> (session 4.3)	887
• Effects of performance-based research funding on publication patterns in the social sciences and humanities – <i>Guns, Raf; Engels, Tim C.E.</i> (session 4.3)	893
• Developing appropriate methods and indicators for evaluation of research in the social sciences and humanities – <i>Sivertsen, Gunnar; Galleron, Ioana</i> (session 5.1)	901
• Clashing Conventions? Exploring Human Resource Management in the Cleavage Between Academic Field Traditions and New Institutional Rules. Quantitative and Qualitative Insights from the Field of Communication and Media Studies in Switzerland – <i>Buhmann, Alexander; Lepori, Benedetto; Ingenhoff, Diana</i> (session 6.3)	904
• A bibliometric indicator with a balanced representation of all fields – <i>Sivertsen, Gunnar</i> (session 6.3)	910
• Measuring research in Humanities and Social Sciences: information from a new Italian data infrastructure – <i>Cicero, Tindaro; Malgarini, Marco</i> (session 6.3)	915
• Trends and Developments in Multi-Authorship in five Social Science disciplines from 1991 to 2014 – <i>Mayer, Sabrina</i> (session 6.3)	924
• ERIH PLUS – Making the SSH visible, searchable and available – <i>Lavik, Gry Ane; Sivertsen, Gunnar</i> (session 7.3)	934
• Indexed University presses: overlap and geographical distribution in five book assessment databases – <i>Mañana-Rodríguez, Jorge; Giménez-Toledo, Elea</i> (session 7.3)	940
• East-African Social Sciences and Humanities Publishing: A Handmade Bibliometrics Approach – <i>Schmidt, Nora</i> (session 7.3)	951
• Alphabetical co-authorship in the social sciences and humanities: evidence from a comprehensive local database – <i>Guns, Raf</i> (session 7.3)	957
• Aligning research assessment in the Humanities to the national Standard Evaluation Protocol. Challenges and developments in the Dutch research landscape – <i>Prins, Ad; Spaapen, Jack; van Vree, Frank</i> (session 8.3)	965

CHAPTER 12: Society, Participation and Culture

• Operationalizing RRI: Relational Quality Assessment & Management Model for Research and Innovation Networks (REQUANET) – <i>Barrenechea, Julieta; Ibarra, Andoni</i> (session 4.6)	971
• What knowledge counts? Insights from an action research project using participatory video with grassroots innovation experiences – <i>Boni, Alejandra; Leivas, Monique; Talón, Alba; De La Fuente, Teresa; Pellicer-Sifres, Victoria; Belda-Miquel, Sergio; López-Fogués, Aurora; Arias, Begoña</i> (session 4.6)	981
• A proposal for measurement of science and innovation culture – <i>Okamura, Asako</i> (session 4.6)	990
• Scientific culture in Colombia. A proposal of an indicator system for science technology and innovation – <i>Pardo Martínez, Clara I.; Alfonso, William H.</i> (session 7.5)	996

• How user-innovators can be identified? Evidence collected from the analysis of practices – <i>Fursov, Konstantin; Nefedova, Alena; Thurner, Thomas</i> (session 7.5)	1005
• Assessing youth engagement with a collaborative Index – <i>Marrades, Ramon; Pajevic, Filipa; Dobrescu, Simona</i> (session 7.5)	1011
• Scientific Culture Measures: Challenges and New Perspectives – <i>José Antonio López Cerezo; Muñoz van den Eynde, Ana</i> (session 8.6)	1012

CHAPTER 13: Indicators in Regional and National Contexts

• Indicators of the Knowledge based Society: Comparison between European and Latin American countries – <i>Villavicencio, Daniel; Morales, Alberto; Amaro, Marcela</i> (session 6.2)	1016
• Measuring cross-border regional STI integration – <i>Makkonen, Teemu</i> (session 6.2)	1025
• From emerging country to a leading role in the scientific and technological field? Analysis of the internationalization of Brazil – <i>Souza, Claudia; De Filippo, Daniela; Sanz-Casado, Elías</i> (session 6.2)	1033
• Measuring internationality without bias against the periphery – <i>Aman, Valeria</i> (session 8.2)	1042
• Indicators on measuring technology convergence worldwide – <i>Luan, Chunjuan</i> (session 8.2)	1050
• Development on the periphery: monitoring science, technology and innovation for sustainable development among Pacific Island Countries – <i>Amaradasa, Ranasinghe; Turpin, Tim</i> (session 8.2)	1059

CHAPTER 14: Networks and Text Analysis

• Breakout discoveries in science: what do they have in common? – <i>Winnink, Jos J.; Tijssen, Robert J.W.; van Raan, Anthony F.J.</i> (session 5.6)	1069
• From university research to innovation - Detecting knowledge transfer via text mining – <i>Woltmann, Sabrina; Clemmensen, Line; Alkærsig, Lars</i> (session 5.6)	1076
• Predicting panel scores by linguistic analysis – <i>Van den Besselaar, Peter; Stout, Luka; Gou, Xiaoli</i> (session 5.6)	1081
• Networks dynamics in the case of emerging technologies – <i>Rotolo, Daniele</i> (session 7.6)	1089
• Using network centrality measures to improve national journal classification lists – <i>Zuccala, Alesia; Robinson-Garcia, Nicolas; Repiso, Rafael; Torres-Salinas, Daniel</i> (session 7.6)	1095
• Bridging centrality: A new indicator to measure the positioning of actors in R&D networks – <i>Scherngell, Thomas; Wanzenböck, Iris; Bergé, Laurent</i> (session 7.6)	1106
• Network heterogeneity in an undirected network – <i>Hu, Xiaojun; Leydesdorff, Loet; Rousseau, Ronald</i> (session 7.6)	1117

CHAPTER 15: Altmetrics

• Article-level metrics and the periphery: an exploration of articles by Brazilian authors – <i>Souza, Iara V.P.; Gouveia, Fabio C.</i> (session 5.7)	1123
• Can we use altmetrics at the institutional level? A case study analysing the coverage by research areas of four Spanish universities – <i>Torres-Salinas, Daniel; Robinson-Garcia, Nicolas; Jiménez-Contreras, Evaristo</i> (session 5.7)	1129
• Enriching the knowledge of altmetrics studies by exploring social media metrics for Economic and Business Studies journals – <i>Nuredini, Kaltrina; Peters, Isabella</i> (session 5.7)	1137
• Comparative study of Colombian Researchers according to data from Google Scholar, ResearchGate and the National System for Measurement Science (Colciencias) – <i>Aguillo, Isidro; Uribe-Tirado, Alejandro; López López, Wilson</i> (session 5.7)	1147
• Next-Generation Metrics: Responsible Metrics & Evaluation for Open Science – <i>Wilsdon, James; Bar-Ilan, Judit; Peters, Isabella; Wouters, Paul</i> (session 6.1)	1156
• A Systematic Identification of Scientists on Twitter – <i>Ke, Qing; Ahn, Yong-Yeol; Sugimoto, Cassidy R.</i> (session 7.7)	1160
• Do Mendeley reader counts reflect the scholarly impact of conference papers? An investigation of Computer Science and Engineering fields – <i>Aduku, Kuku J.; Thelwall, Mike; Kousha, Kayvan</i> (session 7.7)	1165
• Currencies of Science: discussing disciplinary "exchange rates" for citations and Mendeley readership – <i>Costas, Rodrigo; Perianes-Rodriguez, Antonio; Ruiz-Castillo, Javier</i> (session 7.7)	1173
• SSH & the City. A Network Approach for Tracing the Societal Contribution of the Social Sciences and Humanities for Local Development – <i>Robinson-Garcia, Nicolas; van Leeuwen, Thed; Rafols, Ismael</i> (session 7.7)	1183
• Comparing the characteristics of highly cited titles and highly alted titles – <i>Didegah, Fereshteh; Bowman, Timothy D.; Bowman, Sarah; Hartley, James</i> (session 7.8)	1190
• What makes papers visible on social media? An analysis of various document characteristics – <i>Zahedi, Zohreh; Costas, Rodrigo; Larivière, Vincent; Hausteijn, Stefanie</i> (session 7.8)	1196
• Normalization of Mendeley reader impact on the reader- and paper-side – <i>Haunschild, Robin; Bornmann, Lutz</i> (session 7.8)	1200

CHAPTER 16: Citation Impact

• Web of Science coverage and scientific performance of Central and Eastern European countries – <i>Ploszaj, Adam; Olechnicka, Agnieszka</i> † (session 1.7)	1208
• Does size matter? An investigation of how department size and other organizational variables influence on publication productivity and citation impact – <i>Aksnes, Dag W.; Rørstad, Kristoffer; Piro, Fredrik N.</i> † (session 1.7)	1214
• Do usage and scientific collaboration associate with citation impact? – <i>Chi, Pei-Shan; Glänzel, Wolfgang</i> † (session 1.7)	1223
• An approach for the condensed presentation of intuitive citation impact metrics which remain reliable with very few publications – <i>Campbell, David; Tippett, Chantale; Côté, Grégoire; Roberge, Guillaume; Archambault, Éric</i> † (session 2.7)	1229
• A comparison of average-based, percentile rank, and other citation impact indicators – <i>Ruiz-Castillo, Javier; Albarrán, Pedro</i> † (session 2.7)	1241
• How does the scientific progress in developing countries affect bibliometric impact measures of developed countries? A counterfactual case study on China – <i>Stahlschmidt, Stephan; Hinze, Sybille</i> † (session 2.7)	1250
• A comparison of the Web of Science with publication-level classification systems of science – <i>Perianes-Rodriguez, Antonio; Ruiz-Castillo, Javier</i> † (session 3.7)	1257
• Ranking Journals Using Social Choice Theory Methods: A Novel Approach in Bibliometrics – <i>Aleskerov, Fuad T.; Pisyakov, Vladimir; Subochev, Andrey N.</i> † (session 3.7)	1266
• The performance and trend of China's academic disciplines from 2006 to 2014 – <i>Hu, Zhigang; Fu, Huang; Guoqiang, Liang; Xianwen, Wang; DeLong, Tang</i> † (session 3.7)	1274
• Comparing absolute and normalized indicators in scientific collaboration: a study in Environmental Science in Latin America – <i>Cabrini-Grácio, Maria C.; Oliveira, Ely F. T.</i> † (session 3.7)	1282

POSTER SESSION 1: Indicators, Assessment, Funding and Innovation

• The research activity index at the Universitat Politècnica de València (IAP): How an institution can complement national regulation on the productivity of university professors in research and teaching activities. Conejero, J. Alberto; Capilla, José; Sánchez-Ruiz, Luis; Amigó, Vicente; Blasco, Agustín; Botti, Vicent; Cano, Juan; Capmany, José; Chiralt, Amparo; Cuesta, Salomé; García Álvarez-Coque, José María; Jordá, Carmen; Juliá, Juan; Miguel, Pedro; Miranda, Miguel A.; Peris, A.	1292
• Bibliometric indicators and activity scores for digital scholars. Mikki, Susanne; Zygmontowska, Marta	1298
• Mapping scientific controversy in Twitter: the Maya city hoax. Denia, Elena	1301
• Visibility and Impact of Research Data Sets in the Life Sciences supported by a Novel Software Infrastructure. Kramer, Claudia; Jung, Nicole; Tremouilhac, Pierre	1304
• Changes in Scholars' Scientific Knowledge Production Shaped by Bibliometric Measures in Taiwan. Peng, Ming-Te	1307
• Purpose-oriented metrics to assess researcher quality; Duarte, Kedma; Weber, Rosina; Pacheco, Roberto C.S.	1310
• On the relationship between research topics and scientific impact: a study of edible animal research. Castelló-Cogollos, Lourdes; Alexandre-Benavent, Rafael; d'Este, Pablo; Rafols, Ismael	1315
• Evaluation of grants schemes in the context of the national research system based on the publication count and citation data: the grants of the Latvian Council of Science. Kokorevics, Arnis	1318
• New approaches to monitor and evaluate Science, Technology and Innovation in health: a pilot study on the Zika virus. Santos, Paula; Feltrin, Rebeca; Fonseca e Fonseca, Bruna; Barros, Ricardo; Reis, Juliana Gonçalves; Barreto, Jorge; Martins, Fatima; Barreto, Maurício; Lima, Nísia Trindade	1325
• Issues relating to a Brazilian model of graduate courses evaluation: the CAPES system. Vogel, Michely J.M.; Kobashi, Nair Y.	1329
• Performance Based Funding and Researchers' Grant Application Strategies. Johann, David; Neufeld, Jörg	1332
• Impact of research evaluation modes of public research funding on the development of research fields and groups in Estonia. Valdmaa, Kaija; Tõnurist, Piret	1339
• The More Funding Sources, the More Citations? The Feasibility Study of Design on "Funding Diversity Indicator". Chen, Carey Ming-Li	1344
• Accuracy and completeness of funding data in the Web of Science. Álvarez-Bornstein, Belén; Morillo, Fernanda; Bordons, María	1345
• Patent indicators for the Spanish nanotechnology domain. Jürgens, Björn; Herrero-Solana, Víctor	1349
• Best-Practice Benchmarking for Israel: The SNI Scorecard – A Multidimensional Perspective. Maital, Shlomo; Buchnik, Tsipy; Getz, Daphne	1355
• Measuring Global Innovation Activities with Article Visiting Geographical Data. Wang, Xianwen; Fang, Zhichao; Yang, Yang; Wang, Hongyin; Hu, Zhigang	1360
• Public scientists contributing to local literary fiction. An exploratory analysis. Azagra-Caro, Joaquín M.; Fernández-Mesa, Anabel; Robinson-García, Nicolas	1370

• Does collaboration facilitate the performance of enterprise innovation? Lv, Qi; Zhu, Donghua; Huang, Ying; Mitkova, Liliana; Wang, Xuefeng; Ogsuz, Gizem	1374
• Structural Analysis of Redundancy Influence of Local Regions in Renewable Energy R&D Projects in Europe. Laruscain-Sarasola, Jaso; Rio Belver, Rosa María; Garechana, Gaizka	1384
• The discrepancy of patent citation behavior between examiners and inventors: a citation network analysis. Huang, Ying; Zhu, Donghua; Lv, Qi; Porter, Alan L.; Wang, Xuefeng	1389
• How Does Technology Transfer from Universities to Market in China? An Empirical Analysis Based on Invention Patent Assignment. Yang, Yang; Ding, Kun; Zhang, Chunbo; Sun, Xiaoling; Hu, Zhigang	1395
• Large Scale Disambiguation of Scientific References in Patent Databases. Zhao, Kangran; Caron, Emiel; Guner, Stanisław	1405

POSTER SESSION 2: Data Characterisation, Classification, Visualisation and Indicator Design

• Quantifying and visualizing different types of scientific collaboration in Nanoscience and Nanotechnology field. Chinchilla-Rodríguez, Zaida; Miguel, Sandra; Perianes-Rodríguez, Antonio; Olmeda-Gómez, Carlos; Ovalle-Perandones, María-Antonia	1412
• Internal Migration of Scientists in Russia and the USA: the Case of Physicists. Dyachenko, Ekaterina	1419
• The Global Research Identifier Database GRID – Persistent IDs for the World's Research Organisations. Szomszor, Martin; Mori, Andres	1426
• Differential Effects of Scopus vs. Web of Science on University Rankings: A Case Study of German Universities. Horstmann, Wolfram; Schmidt, Birgit	1429
• On the normalization of citation impact based on the Essential Science Indicators classification of Thomson Reuters. Baranova, Olga; Peris, Alfred	1433
• Rock around the clock? Exploring scholars' downloading patterns. Cameron-Pesant, Sarah; Jansen, Yorrick; Larivière, Vincent	1435
• Research leadership in key fields for emerging and developing countries. González-Alcaide, Gregorio; Huamani, Charles; Park, Jinseo; Ramos, José M.	1440
• Mass Gathering as an emerging field: a co-citation analysis. González-Alcaide, Gregorio; Llorente, Pedro; Ramos, José M.	1444
• Research Activity Classification based on Time Series Bibliometrics. Kawamura, Takahiro; Yamashita, Yasuhiro; Matsumura, Katsuji	1449
• Inclusion of Gender perspective in scientific publications in Energy Efficiency. Mauleón, Elba; De Filippo, Daniela	1454
• Gender-based differences in German-language publications. Mayer, Sabrina	1458
• Scientific productivity and the impact of neurosurgery scientists in WOS and Mendeley: a gender study. Sotudeh, Hajar; Dehdarirad, Tahereh; Pooladian, Aida	1463
• How is the counting method for a publication or citation indicator chosen? Gauffriau, Marianne	1467
• The occurrence areas of the dependence problem of the h-index. Liu, Chichen; Cai, Sanfa; Liu, Yuxian	1471
• Multivariate bibliometric analysis of scientific production indicators: a taxonomy of countries scientific degree of centrality. Silva, Deise D.; Grácio, Maria C. C.	1475
• Uncriticized citation process of the indicators like social contagion – a case study of the success rate of commercialization of the public R&D in South Korea. Park, Jinseo; Kim, Sun-Woo; Lee, June Young; Song, Tae Ho	1478
• A comparative analysis of Western Europe and Latin America based on social and scientific indicators. Castanha, Renata; Grácio, Maria C. C.	1482
• Indicators of endogamy and reciprocity in PhD theses assessments. Castelló-Cogollos, Lourdes; Aleixandre Benavent, Rafael; Castelló-Cogollos, Rafael	1487
• Scientific Impact Indicators: a comparative study of Brazilian journals' impact factors. Almeida, Catia C.; Grácio, Maria C. C.	1491
• Sub-fields of Library and Information Science in Turkey: A Visualization Study. Taşkın, Zehra; Doğan, Güleda; Al, Umut	1494
• Content words as measure of structure in the science space. Lamers, Wout S.	1498
• Study on the International and Domestic Subject Areas' Distributions. Wenjie, Wei; Junlan, Yao; Liu, Yuxian	1502
• Characteristics of Paper Publication by Major Countries Focusing on Journals. Fukuzawa, Naomi	1506
• 4D Specialty Approximation: Ability to Distinguish between Related Specialties. Rons, Nadine	1511
• Analysis of Structure of Scientific Publications at Universities Focusing on Sub-Organizations. Murakami, Akiyoshi; Saka, Ayaka; Igami, Masatsura	1514

INDEX BY AUTHOR.....1518

INTRODUCTION

This is the book of proceedings of the 21st Science and Technology Indicators Conference that took place in València (Spain) from 14th to 16th of September 2016.

The conference theme for this year, 'Peripheries, frontiers and beyond' aimed to study the development and use of Science, Technology and Innovation indicators in spaces that have not been the focus of current indicator development, for example, in the Global South, or the Social Sciences and Humanities.

The exploration to the margins and beyond proposed by the theme has brought to the STI Conference an interesting array of new contributors from a variety of fields and geographies.

This year's conference had a record 382 registered participants from 40 different countries, including 23 European, 9 American, 4 Asia-Pacific, 4 Africa and Near East. About 26% of participants came from outside of Europe.

There were also many participants (17%) from organisations outside academia including governments (8%), businesses (5%), foundations (2%) and international organisations (2%). This is particularly important in a field that is practice-oriented.

The chapters of the proceedings attest to the breadth of issues discussed. Infrastructure, benchmarking and use of innovation indicators, societal impact and mission oriented-research, mobility and careers, social sciences and the humanities, participation and culture, gender, and altmetrics, among others.

We hope that the diversity of this Conference has fostered productive dialogues and synergistic ideas and made a contribution, small as it may be, to the development and use of indicators that, being more inclusive, will foster a more inclusive and fair world.

The organising committee

Jordi Molas-Gallart, Alejandra Boni, Elena Castro-Martínez, Ismael Rafols and Richard Woolley

PERIPHERIES, FRONTIERS AND BEYOND

This conference aims to stimulate reflection on the **challenges posed to S&T indicator development and use in geographical, cognitive or social spaces that are peripheral** or marginal to the centres of economic, scientific or technological activity. The focus is also on emerging areas of research and innovation that are inadequately described by existing, quantitative or qualitative indicators.

We propose to **identify, describe and analyse the problems that emerge in situations and spaces where indicators are used beyond their scope of validity**. The conference aims to offer an international platform to propose, and discuss, alternative approaches and indicators.

The conference will consider both weak (technical) and strong (socio-political) notions of periphery. The **weak notion understands peripheries as areas that are not adequately covered or targeted by current indicators**. The main concern here is the existence of indicator biases; the challenge lies in developing approaches and indicators that provide a more accurate or valid representation of science, technology and innovation activities.

The strong notion sees the periphery as composed by those having a lower status in an unequal or dependent relationship. It is therefore a relational concept in a situation that involves structural unequal access to resources. According to this view, peripheries tend to remain as such unless determined efforts to change their situation are undertaken and the use of indicators may contribute to build and sustain peripheral situations. The strong notion of periphery underlines the performative nature of indicators; that is, their capacity to shape reality.

The conference will consider various types of peripheral spaces. In the global economy, some geographical regions are often conceived as peripheral. Developing countries were long ago described as “the” periphery, but within every **geographical territory** we can also encounter peripheral zones (Southern European and Eastern European countries as peripheral to the European Union, poor regions are peripheral to the capital and richer regions within a country, etcetera). Specific problems also emerge in regions that undergo **socio-economic transitions** and are in need of implementing alternative (re)development strategies, in particular in relation to **sustainability**.

We can also refer to peripheral **social groups**: the disenfranchised, the poor, or perhaps the elderly. Research and innovation conducted in these spaces may require different types of indicators from the ones we are accustomed to use. There are also **cognitive peripheries**: areas of research that do not capture the attention of mainstream politicians and receive more limited resources. For example, many fields in the humanities could be considered a peripheral when compared to the mainstream natural sciences or engineering.

Each of these peripheries has their own knowledge generation and application systems and may be better analysed using **tailored indicators, some of which can be of a qualitative rather than quantitative nature**. However, analysts often face resource limitations to develop indicators tailored to the peculiarities of their context and are confronted with the potential use of conventional indicators –which are not fully suited to reflect these contexts. The use of such **indicators may result in inadequate analysis and unintended effects**.

The conference aims to be a platform to **reflect on the potential causes and effects of indicators usage in peripheral spaces**: in mobility and internationalisation, reduction of thematic diversity and alignment or misalignment with local societal needs.

COMMITTEES

CONFERENCE CHAIR

Jordi Molas-Gallart
Ingenio, CSIC-UPV, València

PROGRAMME BOARD

- Judith Sutz · Univ. de la República, Montevideo
- Tim Turpin · Univ. Western Sydney
- Sybille Hinze · DZHW, Berlin
- Robert Tijssen · CWTS, Univ. Leiden
- Jordi Molas-Gallart, Ingenio, CSIC-UPV, València
- Ismael Rafols · Ingenio, CSIC-UPV, València
- Elena Castro-Martínez · Ingenio, CSIC-UPV, València

LOCAL ORGANIZING COMMITTEE

- Alejandra Boni
- Elena Castro-Martínez
- Jordi Molas-Gallart
- Ismael Rafols
- Richard Woolley

SCIENTIFIC COMMITTEE

- Isidro F Aguillo · *IPP-CSIC, Spain*
- Juan Pablo Alperin · *Simon Fraser University, Canada*
- Rigas Arvanitis · *IFRIS, IRD, France*
- Joaquín M. Azagra · *INGENIO (CSIC-UPV), Spain*
- Rodolfo Barrere · *RICYT and Centro Redes, Argentina*
- Julieta Barrenechea · *Euskampus, Universidad del País Vasco, Spain*
- Katy Börner · *School of Informatics and Computing, Indiana University Bloomington, USA*
- Carolina Cañibano · *INGENIO (CSIC-UPV), Spain*
- Elena Castro-Martínez · *INGENIO (CSIC-UPV), Spain*
- Rodrigo Costas · *(CWTS) Leiden University, The Netherlands*
- Susan Cozzens, *Georgia Institute of Technology, Atlanta · USA*
- Laura Cruz Castro · *IPP, CSIC, Spain*
- Pablo D'Este · *INGENIO (CSIC-UPV), Spain*
- Gabriela Dutrénit · *Departamento de Producción Económica, Universidad Autónoma Metropolitana, México*
- Ernesto Fernández Polcuch · *Division of Science Policy and Capacity Building Natural Sciences Sector. UNESCO*
- Sergio Filho-Salles · *Departamento de Política Científica e Tecnológica, State University of Campinas, Brazil*
- Fernando Galindo-Rueda · *OECD Directorate for Science, Technology and Industry, France*
- Fred Gault · *UNU-MERIT, Maastricht University, The Netherlands*
- Elea Giménez Toledo · *Instituto de Filosofía, CSIC, Spain*
- Wolfgang Glanzel, *ECOOM – KU Leuven, Belgium*
- Jochen Gläser · *Center for Technology and Society TU Berlin, Germany*
- Stefanie Haustein, *(EBSI) Université de Montréal, Canada*
- Sybille Hinze · *DZHW, Berlin, Germany*
- Hugo Hollanders · *UNU-MERIT, Maastricht University, The Netherlands*
- Wolfram Horstmann · *University of Göttingen, Germany*
- Erika Kraemer-Mbula · *IERI, Tshwane University of Technology, South Africa*
- Pablo Kreimer · *Universidad de Quilmes, Argentina*
- Frédérique Lang · *SPRU, University of Sussex, UK*
- Philippe Laredo · *IFRIS, Université Paris-Est, France*
- Vincent Larivière · *School of Library and Information Science, Université de Montreal, Canada*
- Thed van Leeuwen · *CWTS, Leiden University, The Netherlands*
- Loet Leydesdorff · *University of Amsterdam, The Netherlands*

COMMITTEES

- Sami Mahroum · *INSEAD, France*
- Luciana Marins · *UNESCO Institute of Statistics*
- Ramon Marrades · *Econcult, University of València, Spain*
- Catalina Martínez-García, *IPP, CSIC. Spain*
- Juan Mateos Garcia · *NESTA, UK*
- Niels Mejlgaard · *Danish Centre for Studies in Research and Research Policy, Aarhus University, Denmark*
- Julia Melkers · *Georgia Institute of Technology, USA*
- Jordi Molas-Gallart · *INGENIO (CSIC-UPV), Spain*
- Johan Mouton · *CREST, Stellenbosch University, South Africa*
- Cameron Neylon · *Curtin University, Perth, Australia*
- Ed Noyons · *(CWTS), Leiden University, The Netherlands*
- Julia Olmos · *Departament Direcció d'Empreses, University of València, Spain*
- Regina Pinto de Gusmao · *Brazilian Ministry of Science, Technology and Innovation (MCTI), Brazil*
- Guilherme Ary Plonski · *Faculdade de Economia, Administração e Contabilidade (FEA), Universidade de Sao Paulo, Brasil*
- Wolfgang Polt · *Joanneum Research, Austria*
- Alan Porter · *Georgia Tech, USA*
- Slavo Radošević · *University College London, UK*
- Ismael Rafols · *INGENIO (CSIC-UPV), Spain*
- Emanuela Reale · *IRCRES, CNR, Italy*
- Nicolas RobinsonGarcia · *INGENIO (CSIC-UPV), Spain*
- Daniele Rotolo · *SPRU, University of Sussex, UK*
- Ronald Rousseau, *KU Leuven, Belgium*
- Mónica Salazar · *InterAmerican Development Bank*
- Mabel Sánchez Barrioluengo · *ISPRA, European Commission, Italy*
- Thomas Scherngell · *Austrian Institute of Technology, Austria*
- Gunnar Siversten · *NIFU, Norway*
- Cassidy Sugimoto · *School of Informatics and Computing, Indiana University Bloomington, USA*
- Judith Sutz · *Universidad República, Uruguay*
- Robert Tijssen · *CWTS, Univ. Leiden, The Netherlands*
- Tim Turpin · *Univ. Western Sydney, Australia*
- Hebe Vessuri · *Instituto Venezolano de Investigaciones Científicas, Venezuela*
- Melanie Walker · *University Free State, South Africa*
- Matthew Wallace · *Canada Foundation for Innovation, Canada*
- Ludo Waltman *Leiden University, The Netherlands*
- Lili Wang · *UNU-MERIT, Maastricht University, The Netherlands*
- Inge van der Weijden · *Leiden University, The Netherlands*
- Richard Woolley · *INGENIO (CSIC-UPV), Spain*
- Paul Wouters · *CWTS, Leiden University, The Netherlands*
- Masaru Yarime · *Graduate School of Public Policy, University of Tokyo, Japan*
- Jon Mikel Zabala Iturriagoitia · *University of Deusto, Spain*
- Zohreh Zahedi · *Leiden University, The Netherlands*
- Ping Zhou · *Department of Information Resources Management. Zhejiang University, China*
- Alesia Ann Zuccala · *The Royal School of Library and Information Science. University of Copenhagen, Denmark*

ORGANIZERS & SPONSORS

ORGANIZERS



PLATINUM SPONSORS



GOLD SPONSOR



SILVER SPONSOR



CONFERENCE PROGRAMME

08.30-13.00 Registration

9.00-10.30 **OPENING SESSION AND 1ST PLENARY SESSION**

Wellcome to STI2016 - *Jordi Molas-Gallart* INGENIO (CSIC-UPV)

Wellcome to the Universitat Politècnica de València - *Francisco J. Mora Mas*, Rector

Keynote lecture: "The Power of Numbers: a critical review of MDG targets for Human Development" *Prof. Sakiko Fukuda-Parr* - The New School, New York, US

10.30-11.00 Coffee Break

11.00-12.30 **PARALLEL SESSIONS 1**

1.1 SPECIAL TRACK 6 · Smart use of indicators for innovation policy

Chair: *Hugo Hollanders & Lili Wang*

Introduction – *Hugo Hollanders*

Innovation indicators: Towards a User's guide – *Michiko Iizuka, Hugo Hollanders*

Analyzing innovation policy indicators through a functional approach: the aeronautic industry case – *Carolina Resende Haddad, Mauricio Maldonado Uriona*

Assessing the performance of national innovation systems in Europe – *Jon Mikel Zabala-Iturriagagoitia*

1.2 Indicators, evaluation and policy

Chair: *Jesper Schneider*

Outlining an analytical framework for mapping research evaluation landscapes – *Fredrik Åström*

When the Brightest are not the Best – *Marco Valent*

The use of indicators and other evidence in two investment decisions of technology innovation – *Nuno F.F.G. Boavida*

1.3 Reward systems

The reward (eco)system of science: More than the sum of its parts? – A Special Fishbowl session – *Nadine Desrochers, Stefanie Haustein, Juan Pablo Alperin, Timothy D. Bowman, Adrián A. Díaz-Faes, Vincent Larivière, Philippe Mongeon, Adèle Paul-Hus, Anabel Quan-Haase, Elise Smith*

1.4 Knowledge Exchange

Chair: *Jan Youtie*

Knowledge integration through collaboration: building indicators using the Diversity/coherence and Proximity frameworks – *Frédérique Lang, Ismael Rafols, Michael Hopkins*

Using a network-based approach to identify interactions structure for innovation in a low-technology intensive sector – *Camille Aouinait*

Interdisciplinary and transdisciplinary institutions: do they constitute peripheries among cultures? – *Bianca Vienni, Ulli Vilsmaier*

“Putting in more than you take out”. Towards evaluating research based on its public (not private) contributions – *Paul Benneworth, Julia Olmos Peñuela, Elena Castro Martínez*

1.5 Careers and labour market

Chair: *Peter van den Besselaar*

Survey on the Labour Market Position of PhD Graduates – *Julia Heuritsch, Cathelijn Waaijer, Inge van der Weijden*

Beyond the indicators: formulation of the career strategies of scientists – *Eva Palinko*

Exploring predictors of scientific performance with decision tree analysis: The case of research excellence in early career mathematics – *Jonas Lindahl*

Stability and longevity in the publication careers of U.S. doctorate recipients – *Cathelijn Waaijer, Benoît Macaluso, Cassidy Sugimoto, Vincent Larivière*

1.6 GENDER SPECIAL SESSION on Gender in science: a periphery?

Chair: *Inge van der Weijden*

Gender equality and evaluation: do fields of science matter? – *Emanuela Reale, Antonio Zinilli*

Scientific and technological output of women and men – *Rainer Frietsch, Susanne Bühner, Patricia Helmich*

Gender and International Mobility of European Researchers – *Carolina Cañibano, Mary Frank Fox and F. Javier Otamendi*

Gender differences and the role of research grants – *Carter Bloch, Evanthia K. Schmidt*

Gender structured universities and their impact on mental health: a focus on PhD students in Flanders – *Katia Levecque & Frederik Anseel*

Gender differences in careers after receiving a personal grant – *Inge van der Weijden & Ingeborg Meijer*

1.7 Citation Impact

Chair: *Rogério Mugnaini*

Determinants of citation impact: A comparative analysis of the Global South versus the Global North – *Hugo Confraria*

PROGRAMME

SEPTEMBER 14

Web of science coverage and scientific performance of Central and Eastern European countries – *Adam Ploszaj, Agnieszka Olechnicka*

Does size matter? An investigation of how department size and other organizational variables influence on publication productivity and citation impact – *Dag W Aksnes, Kristoffer Rørstad, Fredrik N Piro*

Do usage and scientific collaboration associate with citation impact? – *Pei-Shan Chi, Wolfgang Glänzel*

12:30-14:30 **POSTER SESSION 1** (see Appendix 1)

13:00-14:30 Lunch

14:30-16:00 **PARALLEL SESSIONS 2**

2.1 Resource distribution and research contents

Chair: *Jochen Gläser*

Unveiling Research Agendas: a study of the influences on research problem selection among academic researchers – *Mariela Bianco, Judith Sutz*

"If we come out with the wrong answer that may affect investments": Exploring how evaluators were influenced by political considerations during the assessment of societal impact – *Gabrielle Samuel, Gemma Elizabeth Derrick*

Must Metrics Serve the Audit Society? Addressing Marketization in Open Access Publishing and Humanities Analytics – *Christopher Newfield, Christopher Muellerleile*

2.2 SPECIAL SESSION on multiplying methods in the field of research evaluation

Chair: *Inge van der Weijden*

Introduction & recap: *Gemma Derrick*

Provocation: *Paul Wouters*

Demonstrations in pairs:

Wolfgang Kaltenbrunner & Michael Ochsner

Ingeborg Meijer & Carolina Cañibano

Rogério Mugnaini & Nadine Desrochers

Next steps: *Gemma Derrick, Jordi Molas-Gallart and Sarah de Rijcke + Irene Ramos-Vielba*

2.3 SPECIAL TRACK 6 · Smart use of indicators for innovation policy

Chair: *Hugo Hollanders & Lili Wang*

Evidence-based policy learning: the case of the Research Excellence Indicator
Sjoerd Hardeman, Daniel Vertesy

Who sets up the bridge? Tracking scientific collaborations between China and the European Union – *Lili Wang*

A case study about the Colombian Observatory of Science and Technology: between context relevant and internationally comparable indicators – *Mónica Salazar*

2.4 SPECIAL TRACK 1 · Data infrastructure and data quality for evolving research metrics

Chair: *Chris Keene*

Introduction – *Chris Keene*

On the Peripheries of Scholarly Infrastructure: A look at the Journals Using Open Journal Systems – *Juan Pablo Alperin, Kevin Stranack, Alex Garnett*

Why researchers publish in journals not indexed in mainstream databases: training, bridging and gap-filling – *Diego Chavarro, Puay Tang, Ismael Rafols*

Identifying Sources of Scientific Knowledge: classifying non-source items in the WoS – *Clara Calero Medina*

2.5 Careers and labour market

Chair: *Pablo D'Este*

Developing research career indicators using open data: the RISIS infrastructure – *Carolina Cañibano, Richard Woolley, Eric Iversen, Sybille Hinze, Stefan Hornbostel, Jakob Tesch*

On the extent of researcher mobility and indicators for mobility – *Stina Gerdes Barriere*

Progress on mobility and instability of research personnel in Japan: scientometrics on a job-posting database for monitoring the academic job market – *Hirota Kawashima, Yasuhiro Yamashita*

National and international scientific elites: an analysis of Chinese scholars *Fei Shu, Vincent Larivière, Charles-Antoine Julien*

2.6 Gender

Chair: *Monica Gaughan*

What drives the gender gap in STEM? The SAGA Science, Technology and Innovation Gender Objectives List (STI GOL) as a new approach to linking indicators to STI policies – *Ernesto Fernández Polcuch, Martin Schaaper, Alessandro Bello*

Picking the best publications to showcase graduate courses: Do institutional mechanisms reinforce gender differences? – *Jacqueline Leta, Guillaume Cabanac*

What factors influence scientific and technological output: A comparison of Thailand and Malaysia – *Catherine Beaudry, Carl St-Pierre*

PROGRAMME

SEPTEMBER 14

2.7 Citation Impact

Chair: *Wolfgang Glänzel*

An approach for the condensed presentation of intuitive citation impact metrics which remain reliable with very few publications – *David Campbell, Chantale Tippett, Gregoire Cote, Guillaume Roberge, Eric Archambault*

A comparison of average-based, percentile rank, and other citation impact indicators – *Pedro Albarrán, Javier Ruiz-Castillo*

How does the scientific progress in developing countries affect bibliometric impact measures of developed countries? A counterfactual case study on China – *Stephan Stahlschmidt, Sybille Hinze*

The returns to scientific specialization – *Orion Penner, Gaétan de Rassenfosse*

16:00-16:30 Coffee Break

16:30-18:00 **2ND PLENARY SESSION** [Sponsored by RISIS]

Roundtable: **Infrastructures for Inclusive and Open Science and RISIS presentation**

Chair: *Ismael Rafols*

Éric Archambault, Science-Metrix, Montréal, Canada

Chris Keene, JISC, UK

Valentin Bogorov, Thomson-Reuters, Moscow, Russia

Abel Packer Scielo, Sao Paulo, Brazil

Hebe Vessuri, IVIC, Venezuela

Emanuela Reale, IRCRES, CNR, Italy

18:30-19:00 Transfer to the cocktail site

19:00-21:00 **WELCOME COCKTAIL** [SANT MIQUEL DELS REIS]



09.00-10.30 PARALLEL SESSIONS 3

3.1 SPECIAL TRACK 3 · Measuring diverse research “qualities”: indicators of societal impact, engagement, participation, and local relevance

Chair: *Judith Sutz*

Introduction – *Judith Sutz*

‘Productive interactions’ for societal impact: developing a research information system for agriculture (RIS-Agric) at Stellenbosch University, South Africa – *Nelius Boshoff, Harrie Esterhuysen*

Publication patterns in research underpinning impact in REF2014 – *Jonathan Adams*

3.2 University-Industry relations

Chair: *Puay Tang*

Measuring macro-level effects of the global economic recession on university-industry research cooperation – *Joaquín M. Azagra-Caro*

3.3 SPECIAL TRACK 5 · Social sciences and the humanities

Chair: *Theo van Leeuwen*

Introduction – *Theo N. van Leeuwen*

Indicators for research performance in the humanities? The scholars’ view on research quality and indicators – *Michael Ochsner, Sven E. Hug*

Quality criteria and indicators for research in theology – What to do with quantitative measures? – *Silvia Martens, Wolfgang Schatz*

3.4 SPECIAL TRACK 1 · Data infrastructure and data quality for evolving research metrics

Chair: *Chris Keene*

Data quality and consistency in scopus and Web of science in their indexing of Czech Journals – *Pavel Mika, Jakub Szarzec, Gunnar Sivertsen*

Missing citations due to exact reference matching: analysis of a random sample from WoS. Are publications from peripheral countries disadvantaged? *Paul Donner*

Funding acknowledgements in the Web of science: inconsistencies in data collection and standardization of funding organizations – *Jeroen van Honk, Rodrigo Costas, Clara Calero-Medina*

Open data in global environmental change: findings from the community – *Birgit Schmidt*

3.5 SPECIAL TRACK 2 · International benchmarking of innovation: challenges and adequacy for developing and developed regions

Chair: *Luciana Marins*

Introduction – *Luciana Marins*

The impact of methodology in innovation measurement – *Espen Solberg, Lars Wilhelmsen, Markus Bugge*

A critical assessment of the quality and validity of composite indicators of innovation – *Daniel Vertesy*

Innovation strategies in Latin American firms – *Fernando Vargas*

3.6 Gender

Chair: *Jacqueline Leta*

Identifying the gender dimension in research content – *Chantale Tippett, David Campbell, Bastien St. Louis Lalonde, Eric Archambault, Julie Callaert, Katerina Mantouvalou, Lucy Arora*

Gender differences in synchronous and diachronous self-citations – *Gita Ghiasi, Vincent Larivière, Cassidy Sugimoto*

Mapping the author gender-distribution of disease-specific medical research
Jens Peter Andersen, Jesper Wiborg Schneider, Mathias Wullum Nielsen

Indicators for constructing scientific excellence: "Independence" in the ERC starting grant – *Helene Schiffbaenker, Florian Holzinger*

3.7 Citation Impact

Chair: *Erjia Yan*

A comparison of the Web of science with publication-level classification systems of science – *Antonio Perianes-Rodriguez, Javier Ruiz-Castillo*

Ranking journals using social choice theory methods: a novel approach in bibliometrics – *Fuad Aleskerov, Vladimir Pislyakov, Andrey Subochev*

The performance and trend of China's academic disciplines from 2006 to 2014 – *Zhigang Hu*

Comparing absolute and normalized indicators in scientific collaboration: a study in Environmental Science in Latin America – *Maria Cláudia Cabrini Grácio, Ely Francina Tannuri de Oliveira*

10.30-11.00 Coffee break

11.00-12.30 **PARALLEL SESSIONS 4**

4.1 SPECIAL GLOBELICS SESSION on Lessons learned for priority setting and indicators relevant to the impact of research programmes in Europe and Emerging Economies. An evidence-based debate between the research and the policy-shaping community

Chair: *Yannis Caloghirou, Nicholas Vonortas*

Thirty years of European Collaboration in Research and Development: Policy-driven Research Networking and the presence of new knowledge-intensive entrepreneurial ventures – *Yannis Caloghirou, Aimilia Protogerou and Evangelos Siokas*

STI Indicators for Emerging Economies: Experiences from Chile, Brazil and Peru
Adriana Bin, Sergio Salles-Filho, Ana Maria Carneiro, Nicholas Vonortas, Juan Ernesto Sepulveda Alonso and Paula Felicio Drummond de Castro

Use of indicators for research and policy impact evaluation: evidence from Russia – *Konstantin Fursov and Stanislav Zaichenko*

4.2 SPECIAL TRACK 3 · Measuring diverse research “qualities”: indicators of societal impact, engagement, participation, and local relevance

Chair: *Judith Sutz*

Societal impact metrics for non-patentable research in dentistry – *Diana Hicks, Kim Isett, Julia Melkers, Le Song, Rakshit Trivedi*

The Evolution of Scientific Trajectories in Rice: Mapping the Relation between Research and Societal Priorities – *Tommaso Ciarli, Ismael Rafols*

Research Quality Plus (RQ+) A Holistic Approach to Evaluating Research
Robert McLean, Osvaldo Feinstein

4.3 SPECIAL TRACK 5 · Social sciences and the humanities

Chair: *Theo van Leeuwen*

Social Impact Open Repository (SIOR). Transforming the peripheral space of social impact of research – *Mar Joanpere, Elvira Samano*

Je veux bien, mais me citez-vous? On publication language strategies in an anglicized research landscape – *Nadine Desrochers, Vincent Larivière*

Effects of performance-based research funding on publication patterns in the social sciences and humanities – *Raf Guns, Tim Engels*

4.4 Indicators and infrastructure

Chair: *Eric Archambault*

Examining data access and use in science – *Erjia Yan, Mengnan Zhao*

SMS: a linked open data infrastructure for science and innovation studies

Peter Van den Besselaar, Ali Khalili, Al Idrissou, Antonis Loizou, Stefan Schlobach, Frank Van Harmelen

Data Citation Policies of Data Providers within the scope of Longitudinal Studies in Life Course Research – *Anke Reinhold, Marc Rittberger, Nadine Mahrholz*

Stepping up Information Infrastructures and Statistical Reporting-Monitoring the German Excellence Initiative – *Anke Reinhardt*

4.5 SPECIAL TRACK 2 · International benchmarking of innovation: challenges and adequacy for developing and developed regions

Chair: *Luciana Marins*

Innovation dynamics of Salvadoran agro-food industry from an evolutionary perspective – *Elías Humberto Peraza Castaneda, Guillermo Aleixandre Mendizábal*

Elucidate Innovation Performance of Technology-driven Mergers and Acquisitions – *Lu Huang, Kangrui Wang, Huizhu Yu, Lining Shang, Liliana Mitkova*

4.6 Society, participation and culture

Chair: *Julia Melkers*

Operationalizing RRI: Relational Quality Assessment & Management Model for Research and Innovation Networks (REQUANET) – *Julieta Barrenechea, Andoni Ibarra*

What knowledge counts? Insights from an action research project using participatory video with grassroots innovation experiences – *Alejandra Boni, Monique Leivas, Alba Talón, Teresa De la Fuente, Victoria Pellicer-Sifres, Sergio Belda-Miquel, Aurora López-Fogués, Begoña Arias*

A proposal for measurement of science and innovation culture – *Asako Okamura*

4.7 Individual Performance

Chair: *María Bordons*

Information sources – information targets: evaluative aspects of the scientists' publication strategies – *Wolfgang Glanzel, Pei-Shan Chi, Christian Gumpenberger, Juan Gorraiz*

The Effect of Holding a Research Chair on Scientists' Impact – *Seyed Reza Mirnezami, Catherine Beaudry*

Public-private collaboration and scientific impact: an analysis at the level of the individual researcher – *Carter Bloch, Thomas K. Ryan, Jens Peter Andersen*

4.8 Funding and EU collaboration

Chair: *Philippe Larédo*

Examining to What Extent Does the Source of Funding Matter for Scientific Impact. A Case Study of Danish EU FP7 Funded Projects – *Thomas Kjeldager Ryan, Jesper Wiborg Schneider*

The Determinants of National Funding in Trans-national Joint Research: Exploring the Proximity Dimensions – *Emanuela Reale, Andrea Orazio Spinello, Antonio Zinilli*

Beyond funding: What can acknowledgements reveal about credit distribution in science? – *Adèle Paul-Hus, Adrián A. Díaz-Faes, Nadine Desrochers, Rodrigo Costas, Maxime Sainte-Marie, Benoît Macaluso and Vincent Larivière*

Allocating organisational level funding on the basis of Research Performance Based assessments, a comparative analysis of the EU Member States in international perspective – *Koen Jonkers, Thomas Zacharewicz, Benedetto Lepori, Emanuela Reale*

12.30-13.30 **3RD PLENARY SESSION** [Sponsored by IFRIS]

Roundtable: **Global collaboration networks: flat world or centre-periphery structure?**

Chair: *Richard Woolley*

Jonathan Adams, Digital Science, London, UK

Rigas Arvanitis, Director of IFRIS, IRD, Paris, France

Sami Mahroum, INSEAD Innovation and Policy Initiative, Abu Dhabi, United Arab Emirates

Mónica Salazar, InterAmerican Development Bank, Bogota, Colombia

13.30-14.30 Lunch

14:30-16:00 **PARALLEL SESSIONS 5**

5.1 **SPECIAL TRACK 5** · Social sciences and the humanities

Chair: *Theo van Leeuwen*

Developing appropriate methods and indicators for evaluation of research in the social sciences and humanities. Presentation of a new COST Action
Gunnar Sivertsen, Ioana Galleron

5.2 SPECIAL TRACK 3 · Measuring diverse research “qualities”: indicators of societal impact, engagement, participation, and local relevance

Chair: *Judith Sutz*

Impact of Research on Development in Cameroon: convergence between supply and research needs in the food sector – *Minkoua Jules René, Ludovic Temple*

Monitoring the Evolution and Benefits of Responsible Research and Innovation (MoRRI) – a preliminary framework for measuring RRI dimensions – *Niels Mejlgaard, Susanne Buehrer, Erich Griessler, Ralf Lindner, Nikos Maroulis, Ingeborg Meijer, Viola Peter, Ismael Rafols, Tine Ravn, Jack Stilgoe, Lena Tsiouri, Richard Woolley, Angela Wroblewski*

“All this grassroots, real life knowledge”: Comparing perceived with realised concerns of including non-academic evaluators in societal impact assessment – *Gemma Derrick, Gabrielle Samuel*

5.3 SPECIAL SESSION on Predicting STEM Career Success by STI Knowledge Utilization Patterns

Chair: *Barry Bozeman, Jan Youtie*

Career Impacts of Cosmopolitan Collaboration – *Barry Bozeman, Monica Gaughan*

Bounded Collaboration and Changing Core-Periphery Relationships in Sino-Russian Scientific Co-Production – *Abdullah Gök, Maria Karaulova, Philip Shapira*

Going home: why do non-US citizens with US Ph.D. degrees return home?
Stuart Bretschneider

The credibility of policy reporting across learning disciplines – *Jan Youtie*

5.4 SPECIAL SESSION on Performance indicators for areas of innovation: international perspective

Chair: *Guilherme Ary Plonski*

A case study of Be'er Sheva Advanced Technology Park (ATP) in Israel – *Daphne Getz, Eliezer Shein*

Porto Digital: an area of innovation as a lever to transform Recife in Brazil
Guilherme Ary Plonski, Désirée M. Zouain

The influence of Science and Technology parks in Spain – *Andres Barge-Gil, Aurelia Modrego Rico*

5.5 Mission Oriented Research Health

Chair: *Sandro Mendonça*

Using novel computer-assisted linguistic analysis techniques to assess the timeliness and impact of FP7 Health's research – *Vilius Stanciauskas*

Professional impact – *Gustaf Nelhans*

Technology push / market pull indicators in healthcare – *Irina Efimenko, Vladimir Khoroshevsky, Ed Noyons, Evgeny Nochevkin*

Mapping the networks of cancer research in Portugal (1990-2015): initial results – *Oriana Rainho Brás, Jean-Philippe Cointet, João Arriscado Nunes, Leonor David, Alberto Cambrosio*

5.6 Text analysis

Chair: *Stefan Hornbostel*

Breakout discoveries in science: what do they have in common? – *Jos Winnink, Robert J.W. Tijssen, Anthony F.J. van Raan*

From university research to innovation – detecting knowledge transfer via text mining – *Sabrina Larissa Woltmann, Line H. Clemmensen, Lars Alkær sig*

Predicting panel scores by linguistic analysis – *Peter Van den Besselaar*

5.7 Altmetrics

Chair: *Stefanie Haustein*

Article-level metrics and the periphery: an exploration of articles by Brazilian authors – *Iara Vidal Pereira de Souza, Fabio Castro Gouveia*

Can we use altmetrics at the institutional level? A case study analysing the coverage by research areas of four Spanish universities – *Daniel Torres-Salinas, Nicolas Robinson-Garcia, Evaristo Jiménez-Contreras*

Enhancing methodology of altmetrics studies by exploring social media metrics for Economic and Business Studies journals – *Kaltrina Nuredini, Isabella Peters*

Comparative study of Colombian Researchers according to data from Google Scholar, ResearchGate and the National System for Measurement Science (Colciencias) – *Isidro F Aguillo, Alejandro Uribe-Tirado, Wilson López*

16:00-16:30 Coffee Break

PROGRAMME

SEPTEMBER 15

16:30-18:00 **PARALLEL SESSIONS 6**

6.1 Altmetrics – PANEL

Roundtable: **Next-generation metrics: responsible metrics & evaluation for open science**

James Wilsdon, Judit Bar-Ilan, Isabella Peters, Paul Wouters

Chair: *James Wilsdon*

6.2 Geography and performance

Chair: *Michael J Kahn*

Indicators of the knowledge based society: Comparison between European and Latin American countries – *Daniel Villavicencio*

Measuring cross-border regional STI integration – *Teemu Makkonen*

From emerging country to a leading role in the scientific and technological field? analysis of the internationalization of Brazil – *Claudia Daniele de Souza, Daniela De Filippo, Elias Sanz Casado*

6.3 SPECIAL TRACK 5 – Social sciences and the humanities

Chair: *Theo van Leeuwen*

Clashing Conventions? Exploring Human Resource Management in the Cleavage Between Academic Field Traditions and New Institutional Rules. Quantitative and Qualitative Insights from the Field of Communication and Media Studies in Switzerland – *Alexander Buhmann*

A bibliometric indicator with a balanced representation of all fields – *Gunnar Sivertsen*

Measuring research in humanities and social sciences: information from a new Italian data infrastructure – *Marco Malgarini, Tindaro Cicero*

Trends and developments in multi-authorship in five social science disciplines from 1991 to 2014 – *Sabrina Jasmin Mayer*

6.4 SPECIAL TRACK 4 · Collaborations, mobility and internationalization

Chair: *Rigas Arvanitis*

Introduction: *Rigas Arvanitis*

Mobility in the academic careers at the Flemish universities – Results from the human resources in research database – *Noëmi Frea Debacker, Karen Vandeveld*

Gatekeeping African studies: A preliminary insight on what do editorial boards indicate about the nature and structure of research brokerage – *Sandro Mendonça*

6.5 Mission Oriented Research – Health

Chair: *Matthew Wallace*

Access to global health research. Prevalence and cost of gold and hybrid open access – *Stefanie Haustein, Elise Smith, Philippe Mongeon, Fei Shu, Vincent Larivière*

Scientific research on diseases: the distinct profile of developed and developing countries – *Alfredo Yegros*

Biodiversity sustainability of phytomedicine research: a 3-dimensions analysis around the North-South divide – *Philippe Gorry*

In Re the academic cartography of sugar sweetened beverages: scientific and technical information, interdisciplinarity, and legal academia – *Lexi C. White*

6.6 Project and programme assessment

Chair: *Diana Hicks*

An assessment of EU-funded research projects: innovators and their innovative potential – *Daniel Nepelski, Vincent Van Roy, Eoghan O'neill*

Evaluating the impact of public space investments with limited time and funds: (methodological) lessons from a Swiss case study – *Franz Barjak*

Researchers and institutions in the periphery: challenges in measuring research capacity for geographically specific programs in the U.S – *Julia Melkers*

Assessing marine biotechnology research centres in peripheral regions: developing global and local STI indicators – *Antoine Schoen, Douglas Robinson*

6.7 Measuring Innovation

Chair: *Joaquín M. Azagra-Caro*

Baseline of indicators for R&D and Innovation in ICT: a tool for decision-making, design and monitoring of public policies – *Henry Mora Holguín, Diana Lucio-Arias, Sandra Zárate, Nayibe Castro, Clara Pardo*

Measuring originality: common patterns of invention in research and technology organizations – *David Li Tang, Erica Wiseman, Tamara Keating, Jean Archambeault*

Linking international trademark databases to inform IP research and policy
Stephen Michael Petrie

Detecting emerging trends and country specializations in energy efficiency – *Daniela De Filippo, Andres Pandiella-Dominique, Elba Mauleon*

PROGRAMME

SEPTEMBER 15

18:00-19:30 ENID General Assembly

18:00-19:30 **SPECIAL SESSION** "Ciencia, Tecnología, Sociedad e Innovación ¿Medimos lo que debemos? ¿Medimos bien?"

Organised and sponsored by the Cátedra CTSi (OEI-Junta de Andalucía)

(This special session is in Spanish)

Chair: Manuel Torralbo Junta de Andalucía, Spain

Judith Sutz – Universidad de la República, Uruguay

Hebe Vessuri – CONICET, Argentina

José Navarrete – Junta de Andalucía, Spain

20:00-22:00 **CONFERENCE DINNER** (Hotel Astoria)



09.00-10.30 **4th PLENARY SESSION**

Conference "The deep structure of STI indicators: Contextual knowledge and scientometrics"

Chair: *Philippe Laredo*

Keynote speaker: *Prof. Johann Mouton* – Stellenbosch University, South Africa

10.30-11.00 Coffee Break

11.00-12.30 **PARALLEL SESSIONS 7**

7.1 Innovation in Government

Roundtable: **SPECIAL SESSION** on Measuring Innovation in Government –

Anthony Arundel, Carter Bloch, Ilka Lakaniemi, Sami Mahroum

Chair: *Sami Mahroum*

7.2 Mission Oriented Research – Agriculture

SPECIAL PANEL on Metrics and Agricultural Science measuring

Multidisciplinary and Applied Research – *Vanessa Méry, Hugo Besemer, Ellen Fest, Soizic Messiaen*

Chair: *Ilkay Holt*

7.3 **SPECIAL TRACK 5** · Social sciences and the humanities

Chair: *Theo van Leeuwen*

ERIH PLUS – Making the SSH visible, searchable and available – *Gry Ane Vikanes Lavik, Gunnar Sivertsen*

Indexed University presses: overlap and geographical distribution in five book assessment databases – *Jorge Mañana-Rodríguez, Elea Giménez-Toledo*

East-African Social Sciences and Humanities Publishing – A Handmade Bibliometrics Approach – *Nora Schmidt*

Alphabetical co-authorship in the social sciences and humanities: evidence from a comprehensive local database – *Raf Guns*

7.4 **SPECIAL TRACK 4** · Collaborations, mobility and internationalization

Chair: *Rigas Arvanitis*

Scientific mobility of Early Career Researchers in Spain and The Netherlands through their publications – *Nicolas Robinson-García, Carolina Cañibano, Richard Woolley, Rodrigo Costas*

The network of international student mobility – *Eva Maria Voegtle, Michael Windzio*
Big Science, co-publication and collaboration: getting to the core – *Michael J Kahn*

Autonomy vs. dependency of scientific collaboration in scientific performance

Zaida Chinchilla-Rodríguez, Sandra Miguel, Antonio Perianes-Rodríguez, María-Antonia Ovalle-Perandones, Carlos Olmeda-Gómez

7.5 Culture and engagement

Chair: *Bianca Vienni*

Scientific culture in Colombia. A proposal of an indicator system for science, technology and innovation – *Clara Pardo, William Alfonso*

How user-innovators can be identified? Evidence collected from the analysis of practices – *Konstantin Fursov*

Assessing youth engagement with a collaborative Index – *Ramón Marrades*

7.6 Networks

Chair: *Ludo Waltman*

Networks dynamics in the case of emerging technologies – *Daniele Rotolo*

Using network centrality measures to improve national journal classification lists – *A. Zuccala, N. Robinson-Garcia, R. Repiso, D. Torres-Salinas.*

Bridging centrality: A new indicator to measure the positioning of actors in R&D networks – *Thomas Scherngell, Laurent Berge, Iris Wanzenböck*

Network heterogeneity in an undirected network – *Xiaojun Hu, Loet Leydesdorff, Ronald Rousseau*

7.7 Altmetrics

Chair: *Juan Pablo Alperín*

A Systematic Identification of Scientists on Twitter – *Qing Ke, Yong-Yeol Ahn, Cassidy R. Sugimoto*

Do Mendeley reader counts reflect the scholarly impact of conference papers? A comparison between ComputerScience and Engineering field. – *Kuku Joseph Aduku, Mike Thelwall, Kayvan Kousha*

Currencies of science: discussing disciplinary “exchange rates” for citations and Mendeley readership – *Rodrigo Costas, Antonio Perianes-Rodríguez, Javier Ruiz-Castillo*

SSH & the City. A Network Approach for Tracing the Societal Contribution of the Social Sciences and Humanities for Local Development – *Nicolas Robinson-Garcia, Thed N. van Leeuwen, Ismael Rafols*

12.30-14.30 POSTER SESSION 2 (see Appendix 2)

13:00-14:30 Lunch

14:30-16:00 PARALLEL SESSIONS 8

8.1 Indicators' use and effects

Chair: *Paul Wouters*

Why DORA does not stand a chance in the biosciences – *Jochen Gläser*

Are institutional missions aligned with journal-based or document-based disciplinary structures? – *Richard Klavans, Kevin Boyack*

Science policy through stimulating scholarly output Does it work? – *Peter Van den Besselaar*

The need for contextualized scientometric analysis: An opinion paper – *Ludo Waltman*

8.2 National systems in the periphery

Chair: *Daniel Villavicencio*

Measuring internationality without bias against the periphery – *Valeria Aman*

Indicators on measuring technology convergence worldwide – *Chunjuan Luan*

Development on the Periphery: monitoring science, technology and innovation for sustainable development among Pacific Island Countries – *Tim Turpin, Ranasinghe Wasantha Amaradasa*

Fake Academic Degrees as an Indicator for Severe Reputation Crisis in the Scientific Community - *Andrey Rostovtsev, Alexander Kostinskiy*

8.3 SPECIAL TRACK 5 · Social sciences and the humanities

Chair: *Theo van Leeuwen*

A SPECIAL DEBATE on Aligning research assessment in the Humanities to the national Standard Evaluation Protocol Challenges and developments in the Dutch research landscape – *Ad Prins, Jack Spaapen, Frank van Vree*

8.4 SPECIAL TRACK 4 · Collaborations, mobility and internationalization

Chair: *Rigas Arvanitis*

The world network of scientific collaborations between cities: domestic or international dynamics? – *Marion Maisonneuve, Denis Eckert, Michel Grossetti, Laurent Jégou, Béatrice Milard*

Trends in the inter-regional and international research collaboration of the PRC's regions: 2000-2015 – *Marc Luwel, Erik van Wijk, Lambertus (Bert) J van der Wurff, Lili Wang*

Iran's scientific dominance and the emergence of South-East Asian countries in the Arab Gulf Region – *Henk F. Moed*

How international is internationally collaborated research? A bibliometric study of Russian surname holder collaboration networks – *Maria Karaulova, Abdullah Gök and Philip Shapira*

8.5 Mission-Oriented Research-Health

Chair: *Tommaso Ciarli*

Network analysis to support research management: evidence from the Fiocruz Observatory in Science, Technology and Innovation in Health – *Bruna de Paula Fonseca e Fonseca Fonseca, Ricardo Barros Sampaio, Marcus Vinicius Pereira da Silva, Paula Xavier dos Santos*

Partial alphabetical authorship in medical research: an exploratory analysis
Philippe Mongeon, Elise Smith, Bruno Joyal, Vincent Larivière

The bibliometric behaviour of an expanding specialisation of medical research
Jonathan Levitt, Mike Thelwall

8.6 SPECIAL SESSION · Scientific Culture Measures. Challenges and New Perspectives

Presentation / Introduction to the topic: What is scientific culture and what is not? – *José Antonio López Cerezo*

What does it mean to be scientifically literate? – *Belén Laspra*

New tools and indicators to measure scientific culture – *Ana Muñoz van den Eynde*

New cultural factors influencing the innovation measures – *María Cornejo Cañamares*

8.7 Altmetrics

Chair: *Rodrigo Costas*

Comparing the characteristics of highly cited titles and highly alted titles
Fereshteh Didegah, Timothy D. Bowman, Sarah Bowman, James Hartley

What makes papers visible on social media? An analysis of various document characteristics – *Zohreh Zahedi, Rodrigo Costas, Vincent Larivière, Stefanie Haustein*

Normalization of Mendeley reader impact on the reader- and paper-side
Robin Haunschild, Lutz Bornmann

16:00-16:30 Coffee Break

16:30-18:00 **5TH PLENARY SESSION** [Sponsored by Thomson Reuters]

Roundtable: **Use of indicators in policy and inclusive metrics**

Chair: *Jordi Molas-Gallart*

Richard Deiss, Directorate General for Research and Innovation, European Commission

Diana Hicks, Georgia Tech, Atlanta, US

Slavo Radosevic, UCL, London, UK

Judith Sutz, President of Globelics & Univ. de la República, Montevideo, Uruguay

18:15-19:30 **STI2016 Fringe**

OPEN SESSION on local examples of participatory research

Video presentations

A roundtable on quality criteria and indicators for Participatory Action Research

Sandra Boni, *Ramon Marrades* and local Valencian activists

18:30-21:30 **CLOSING COCKTAIL & MUSIC**



12:30-14:30 **POSTER SESSION 1**

INDICATORS, ASSESSMENT, FUNDING AND INNOVATION

The research activity index at the Universitat Politècnica de València (IAIP): How an institution can complement national regulation on the productivity of university professors in research and teaching activities. *Conejero, J. Alberto; Capilla, José; Sánchez-Ruiz, Luis; Amigó, Vicente; Blasco, Agustín; Botti, Vicent; Cano, Juan; Capmany, José; Chiralt, Amparo*

Bibliometric indicators and activity scores for digital scholars. *Mikki, Susanne; Zygmuntowska, Marta*

Mapping scientific controversy in Twitter: the Maya city hoax. *Denia, Elena*

Visibility and Impact of Research Data Sets in the Life Sciences supported by a Novel Software Infrastructure. *Kramer, Claudia; Jung, Nicole; Tremouilhac, Pierre*

Changes in Scholars' Scientific Knowledge Production Shaped by Bibliometric Measures in Taiwan. *Peng, Ming-Te*

Purpose-oriented metrics to assess researcher quality; *Duarte, Kedma; Weber, Rosina; Pacheco, Roberto C.S.*

On the relationship between research topics and scientific impact: a study of edible animal research. *Castelló-Cogollos, Lourdes; Aleixandre-Benavent, Rafael; d'Este, Pablo; Rafols, Ismael*

Evaluation of grants schemes in the context of the national research system based on the publication count and citation data: the grants of the Latvian Council of Science. *Kokorevics, Arnis*

New approaches to monitor and evaluate Science, Technology and Innovation in health: a pilot study on the Zika virus. *Santos, Paula; Feltrin, Rebeca; Fonseca e Fonseca, Bruna; Barros, Ricardo; Reis, Juliana Gonçalves; Barreto, Jorge; Martins, Fatima; Barreto, Maurício; Lima, Nísia Trindade*

Issues relating to a Brazilian model of graduate courses evaluation: the CAPES system. *Vogel, Michely J.M.; Kobashi, Nair Y.*

Performance Based Funding and Researchers' Grant Application Strategies. *Johann, David; Neufeld, Jörg*

Impact of research evaluation modes of public research funding on the development of research fields and groups in Estonia. *Valdmaa, Kaija; Tõnurist, Piret*

The More Funding Sources, the More Citations? The Feasibility Study of Design on "Funding Diversity Indicator". *Chen, Carey Ming-Li*

12:30-14:30 **POSTER SESSION 1**

INDICATORS, ASSESSMENT, FUNDING AND INNOVATION

Accuracy and completeness of funding data in the Web of Science. *Álvarez-Bornstein, Belén; Morillo, Fernanda; Bordons, María*

Patent indicators for the Spanish nanotechnology domain. *Jürgens, Björn; Herrero-Solana, Víctor*

Best-Practice Benchmarking for Israel: The SNI Scorecard – A Multidimensional Perspective. *Maital, Shlomo; Buchnik, Tsipy; Getz, Daphne*

Measuring Global Innovation Activities with Article Visiting Geographical Data. *Wang, Xianwen; Fang, Zhichao; Yang, Yang; Wang, Hongyin; Hu, Zhigang*

Public scientists contributing to local literary fiction. An exploratory analysis. *Azagra-Caro, Joaquín M.; Fernández-Mesa, Anabel; Robinson-Garcia, Nicolas*

Does collaboration facilitate the performance of enterprise innovation? *Lv, Qi; Zhu, Donghua; Huang, Ying; Mitkova, Liliana; Wang, Xuefeng; Ogsuz, Gizem*

Structural Analysis of Redundancy Influence of Local Regions in Renewable Energy R&D Projects in Europe. *Larruscain-Sarasola, Jaso; Rio Belver, Rosa María; Garechana, Gaizka*

The discrepancy of patent citation behavior between examiners and inventors: a citation network analysis. *Huang, Ying; Zhu, Donghua; Lv, Qi; Porter, Alan L.; Wang, Xuefeng*

How Does Technology Transfer from Universities to Market in China? An Empirical Analysis Based on Invention Patent Assignment. *Yang, Yang; Ding, Kun; Zhang, Chunbo; Sun, Xiaoling; Hu, Zhigang*

Large Scale Disambiguation of Scientific References in Patent Databases. *Zhao, Kangran; Caron, Emiel; Guner, Stanisław*

12:30-14:30 **POSTER SESSION 2**

DATA CHARACTERISATION, CLASSIFICATION, VISUALISATION AND INDICATOR DESIGN

Quantifying and visualizing different types of scientific collaboration in Nanoscience and Nanotechnology field. *Chinchilla-Rodríguez, Zaida; Miguel, Sandra; Perianes-Rodríguez, Antonio*

Internal Migration of Scientists in Russia and the USA: the Case of Physicists. *Dyachenko, Ekaterina*

The Global Research Identifier Database GRID – Persistent IDs for the World's Research Organisations. *Szomszor, Martin; Mori, Andres*

Differential Effects of Scopus vs. Web of Science on University Rankings: A Case Study of German Universities. *Horstmann, Wolfram; Schmidt, Birgit*

On the normalization of citation impact based on the Essential Science Indicators classification of Thomson Reuters. *Baranova, Olga; Peris, Alfred*

Rock around the clock? Exploring scholars' downloading patterns. *Cameron-Pesant, Sarah; Jansen, Yorrick; Larivière, Vincent*

Research leadership in key fields for emerging and developing countries. *González-Alcaide, Gregorio; Huamaní, Charles; Park, Jinseo*

Mass Gathering as an emerging field: a co-citation analysis. *González-Alcaide, Gregorio; Llorente, Pedro; Ramos, José M.*

Research Activity Classification based on Time Series Bibliometrics. *Kawamura, Takahiro; Yamashita, Yasuhiro*

Inclusion of Gender perspective in scientific publications in Energy Efficiency. *Mauleón, Elba; De Filippo, Daniela*

Gender-based differences in German-language publications. *Mayer, Sabrina*

Scientific productivity and the impact of neurosurgery scientists in WOS and Mendeley: a gender study. *Sotudeh, Hajar; Dehdarirad, Tahereh; Pooladian, Aida*

How is the counting method for a publication or citation indicator chosen? *Gauffriau, Marianne*

The occurrence areas of the dependence problem of the h-index. *Liu, Chichen; Cai, Sanfa; Liu, Yuxian*

Multivariate bibliometric analysis of scientific production indicators: a taxonomy of countries scientific degree of centrality. *Silva, Deise D.; Grácio, Maria C. C.*

12:30-14:30 POSTER SESSION 2

DATA CHARACTERISATION, CLASSIFICATION, VISUALISATION AND INDICATOR DESIGN

Uncriticized citation process of the indicators like social contagion – a case study of the success rate of commercialization of the public R&D in South Korea Park. *Jinseo; Kim, Sun-Woo; Lee, June Young; Song, Tae Ho*

A comparative analysis of Western Europe and Latin America based on social and scientific indicators. *Castanha, Renata; Grácio, Maria C. C.*

Indicators of endogamy and reciprocity in PhD theses assessments. *Castelló-Cogollos, Lourdes; Aleixandre Benavent, Rafael; Castelló-Cogollos, Rafael*

Scientific Impact Indicators: a comparative study of Brazilian journals' impact factors. *Almeida, Catia C.; Grácio, Maria C. C.*

Sub-fields of Library and Information Science in Turkey: A Visualization Study. *Taşkın, Zehra; Doğan, Güleda; Al, Umut*

Content words as measure of structure in the science space. *Lamers, Wout S.*

Study on the International and Domestic Subject Areas' Distributions. *Wenjie, Wei; Junlan, Yao; Liu, Yuxian*

Characteristics of Paper Publication by Major Countries Focusing on Journals. *Fukuzawa, Naomi*

4D Specialty Approximation: Ability to Distinguish between Related Specialties. *Rons, Nadine*

Analysis of Structure of Scientific Publications at Universities Focusing on Sub-Organizations. *Murakami, Akiyoshi; Saka, Ayaka; Igami, Masatsura*



PLENARIES

You can watch the plenaries on youtube:

<https://www.youtube.com/watch?v=uEmouz4TOds&list=PLBAoGA5erbiU9e9njUnLBBC0VhgbUHBzb>

SAKIKO FUKUDA-PARR

THE NEW SCHOOL, NEW YORK, US



Sakiko Fukuda-Parr is a development economist interested in human development and the broad question of national and international policy strategies. She is currently a Professor at The New School, in the International Affairs Program where she chairs the Development Concentration. From 1995 to 2004, Sakiko was lead author and director of the UNDP Human Development Reports. Previously, she worked at the World Bank and UNDP on agriculture, aid coordination in Africa and capacity development. Recently, United Nations Secretary-General Ban Ki-moon announced the appointment of Prof Fukuda-Parr as a member of the newly established high-level panel on health technology and access to medicines.

One of her current research projects is "The Power of Numbers: A Critical Review of MDG Targets for Human Development and Human Rights (co-coordinator with Alicia Yamin, Harvard University) – a multi-author research initiative on the impact of global goal setting on international development agendas".

14 Septembre, 16:30 – 18:00 · Room 1 · ground floor

Roundtable: Infrastructures for Inclusive and Open Science and RISIS presentation**Panelists:** Éric Archambault (Science-Metrix. Montréal, Canada), Valentin Bogorov (Thomson-Reuters. Moscow, Russia), Abel Packer (SciELO, Sao Paulo. Brazil), Hebe Vessuri (IVIC. Venezuela).**Chair:** Ismael Ràfols (INGENIO, CSIC-UPV. València, Spain)

The infrastructure for information on S&T has a strong influence on the patterns of communication and the visibility of science. Scientific journals and the bibliographic database shape the production, circulation and consumption of knowledge. Since the mid 20th century, science dynamics was influenced by Garfield's notion that a small "core journals" that published most of the all the research of significance – those covered by the ISI (now Web of Science) database. These core journals of 'international' scope that 'controlled' most scientific communication were mainly published in a few Western countries. The databases were often used by managers to stratify science into high quality cores (top quartile journals), second class science and 'invisible science'.

Since the 1980s, researchers in the global south and in some disciplines such SSH have increasingly voiced discontent about Garfield notion of 'core', in particular about its consequences in terms of the invisibility of 'peripheral' journals and the effects of journal stratification on knowledge production. For example, there have been worries of suppression of research on topics relevant to developing countries or marginalised populations which are published in local journals in languages other than English.

Also, the great changes in ICT in the last two decades have facilitated the pluralisation of scientific information. The appearance of new databases, such as SciELO or Redalyc that explicitly aim to fill in gaps in coverage. Moreover, the advent of open access technologies that can make 'local' journals accessible across the globe. Also new forms of science dissemination, such as blogs or twitter, or new forms of publishing (e.g. data sharing), are making scientific information more diverse. However, this succession of transformations towards more 'open science' poses major challenges to the governance of information infrastructure.

In this round table we aim to discuss, first, the diverse strategies for developing infrastructure with an open and comprehensive coverage and, second, the governance of the scientific information infrastructure in the face of new forms of communication.

First, current general databases have a limited coverage while more comprehensive databases are specific to some regions or sectors. Thus, most S&T indicators and benchmarking are based on conventional 'core' databases. Should more comprehensive databases be developed, mixing different types of science – e.g. more 'local' and more 'universal'? How should indicators of these databases be interpreted? How is open access best provided and maintained?

Second, the development of robust and publicly trusted indicators needs an open and transparent data infrastructure. What type of governance should be established to ensure public critical analysis? Which types of organisations should manage the data? Should these be distributed or centralised systems?

Previous studies of standards and infrastructure have shown that deep political implications of apparently technical choices. If we aim to make science more open, democratic and inclusive, we need to be highly reflective on how we develop these infrastructures.

15 Septembre, 12:30 – 13:30 · Room 1 · ground floor

Global networks, internationalization and local research agendas: indicators for benchmarking or context specific?

Panelists: Jonathan Adams (Digital Science, London, UK), Rigas Arvanitis (Director of IFRIS, IRD, Paris, France), Sami Mahroum, (INSEAD Innov. and Policy Initiative, Abu Dhabi, United Arab Emirates), Mónica Salazar (InterAmerican Development Bank, Bogotá, Colombia).

Chair: Richard Woolley (Ingenio, CSIC-UPV, València, Spain).



It is widely accepted that 'global science' or the globalization of scientific work, collaboration and coordination has developed rapidly in the era of mass long-haul travel and has intensified with the arrival of the 'Internet age'. The ideal of a global science network through which access and contribution to science is no longer structured by zones of inclusion and exclusion is said to be within reach. In this so-called 'flat-earth' view of globalized science, physical location and local resources are secondary to international networks. Strategies for raising scientific quality are contingent on plugging into the global networks. Through these networks, countries with lower resource levels (human capital, research infrastructure, financial) are expected to access advance knowledge and techniques. This is assumed to lead to a faster rising level of competence underpinning the advancement of a science and innovation driven mode of socio-economic development.

Indicators of 'internationalization' thus become important for monitoring global connectedness as a proxy for a network model of development. Countries that map and understand their collaborations can leverage their strengths and use policy interventions to build global links in targeted areas. Indicators play an important role in highlighting opportunities and progress in connecting to key global channels. Research quality is assumed to rise in concert with internationalization indicators, lifting downstream activities and oppor-

tunities for commercial exploitation. Indicators that seek to benchmark or produce universalized measures (such as the global university rankings) are therefore regarded as relevant and seen as having positive impacts on the direction of policy development.

In contrast to this vision of global equalization, another interpretation of the globalized organization of science sees the global networks as a perpetuation of asymmetric relations of power and control over the scientific agenda. In this view, global networks mainly operate to export the research agenda of the rich and successful countries to distributed research groups in other locations. The development of a science that is not just of high quality but also of relevance to its context may be hampered by focusing on the research questions which are of interest to researchers and funding agencies in highly developed countries.

Indicator development faces other challenges according to this view that the scientific world is very far from being 'flat'. Different types of indicators might be needed in different contexts. 'Universal' measures such as global rankings may be useless, or even potentially misleading, in terms of shaping policy agendas in these contexts.

Taking these polar views, we can see that the same global network could be interpreted in two very different ways. Perhaps the challenge is to find the complementarities between these two visions. Perhaps a more reflexive politics of responsible indicator development is needed. What exactly should be the role of state administrations in this contested terrain, including those charged with capturing and presenting data for S&T information systems? This session will bring these issues of the global and the local/regional into focus and into question. It will provide an opportunity for robust debate and for challenging perspectives on the received vision of 'global science' and the indicators of internationalization that help to construct this vision.

JOHANN MOUTON

STELLENBOSCH UNIVERSITY, SOUTH AFRICA



Johann Mouton is Professor in and Director of the Centre for Research on Evaluation, Science and Technology at Stellenbosch University and the African Doctoral Academy. Johann Mouton is also the Programme Director of five post-graduate programmes in Monitoring and Evaluation Studies and Science and Technology Studies. He has authored or co-authored 10 monographs including *Understanding social research* (1996), *The practice of social research* (2002, with E. Babbie) and *How to succeed in your Masters and doctoral studies* (2001). He has supervised or co-supervised 70 doctoral and master's students. He received two prizes from the Academy for Science and Arts in South Africa including one for his contribution to the promotion of research methodology in South Africa. In 2012 he was elected to the Council of the Academy of Science of South Africa.

His main research interests are the philosophy and methodology of the social sciences, higher education knowledge production, sociology of science, scientometrics and science policy studies.

16 Septembre, 16:30 – 18:00 · Room 1 · ground floor

Roundtable on “Use of indicators in policy and inclusive metrics”

Panellists: Richard Deiss, (Directorate General for Research and Innovation, European Commission), Diana Hicks (Georgia Tech. Atlanta, USA), Slavo Radosevic (UCL. London, UK), Judith Sutz (President of Globelics & Univ. de la República. Montevideo, Uruguay).

Chair: Jordi Molas-Gallart (INGENIO (CSIC-UPV), Spain).



The STI conferences have long aimed to stimulate reflection on the use of indicators. Two years ago, in a plenary roundtable on “quality standards for evaluation indicators” Diana Hicks launched the idea of a “manifesto” that would lay out some basic principles on the evaluative use of indicators. This led to the Leiden Manifesto for Research Metrics, a set of “ten principles to guide research evaluation”. The Leiden Manifesto has become an influential initiative to raise awareness of the challenges posed by the use of indicators in evaluation and, therefore, to inform policy decisions. The HEFCE report *The Metric Tide* also recommended general principles such as robustness, humility, transparency, diversity and reflexivity regarding the responsible use of research metrics. Yet, although these principles have been well received, in many cases they do not provide solutions but state desirable goals. Agreement with the principles does not imply the capacity to implement them. How can we move from general principles to more specific advice?

This closing roundtable will discuss how to address the challenges posed by the use of indicators in policy, in particular in relation to geographical, cognitive or social areas that are not well described by current indicators.

First, we need to consider how indicators are used in the policy process. There is agreement among many evaluation practitioners that “quantitative evaluation should support qualitative, expert assessment”, as stated by the first principle of the Leiden Manifesto. Indicators and the analyses based on indicators should therefore inform but not substitute judgement. How can the principle operate in practice? Is this applicable in all circumstances? Can the application of mixed methods to evaluation help address this problem?

A second challenge relates to the adequacy of currently available indicators for assessing institutions or research against their stated missions and their specific context. The indicators community has developed sensible methods for measuring performance against some missions in certain contexts. However, some fields, such as the Humanities, or missions, such as health care, and many regions, are currently poorly covered by indicators. How can we use indicators to inform policy when they are known to be biased, for example due to the uneven topic or country coverage of databases? How should we use indicators so that local research and innovation is made visible and valued? How can we, for instance, use indicators to capture the performance of an organisation against its research missions when these are peculiar to a local context? What are the opportunities for the development and use of alternative indicators that are inclusive of currently invisible or marginalised research and innovation?

We would like to invite the panellists and the audience to share ideas and collective initiatives so that our community can contribute to a wiser, more inclusive and responsible use of S&T indicators.

ORAL PRESENTATIONS

CHAPTER 1

Data Infrastructure and Data Quality



Must Metrics Serve the Audit Society? Addressing Marketization in Open Access Publishing and Humanities Analytics¹

Christopher Newfield^{*}, Christopher Muellerleile^{**}

^{*}cnewf@english.ucsb.edu

Department of English, University of California, Santa Barbara, CA 93106 (USA)

^{**} c.m.muellerleile@swansea.ac.uk

Department of Geography, Swansea University, Singleton Park, Swansea, SA2 8PP

ABSTRACT

The “audit society” (Power 1997) seemed to arise from its efforts to increase accountability and efficiency in public institutions. Accounting and its numerical indicators was to be a neutral tool to accomplish audits, which in turn were to increase fairness and transparency in the institutions of democratic society.

Since then, two related activities have been proceeding in parallel: the technical and institutional development of indicators, and the analysis of their institutional and sociocultural effects. Our starting point is the distance and frequent divergence of these two activities: while indicators have advanced and proliferated, their impact has been subject to largely negative critique. Most academics are habituated to ubiquitous assessment. And yet complaints about this are equally ubiquitous. Can these practices—numerical assessment and its critique—be brought into a productive relationship?

On the surface, the answer seems to be an obvious yes. Both producers and consumers of indicators release guidelines and standards designed to spread knowledge of the limits of numerical indicators and to reduce abuse (Archambault and Gagné 2004; NICE 2013). Such guidelines invariably call for the embedding of quantitative metrics in the appropriate institutional and professional contexts. For example, the important “Leiden Manifesto” has as its first principle, “Quantitative evaluation should support qualitative, expert assessment” (Hicks et al., 2015). Similarly, the editor of *Times Higher Education*, which offers elaborate university ranking services, insists on the great value of metrics as long as consumers realize that “contextual information is vitally important” (Gill 2015). Most academics agree that valid numerical indicators can be constructed (Gingras 2014) and can be used correctly to assess research impact and productivity in the context of “informed peer review” (Wouters 2014). Core principles for the valid use of metrics are: (1) maintain the specificities and purposes of the evaluative context; (2) link quantitative to qualitative analysis; (3) include

¹ This work was supported by the Universities in the Knowledge Economy (European Commission Initial Training Network) and by the Limits of the Numerical research project (Universities of Cambridge, Chicago, and California at Santa Barbara).

professional expertise and substantive domain knowledge; (4) use data and process transparency to bind evaluators and evaluatees in a shared community. And yet when institutional analysts look at the impact of numerical assessment over time, do they find these principles in action?

This paper argues that the root problem with quantitative metrics is marketization as we define it here—that although quantitative information can be used in a wide range of ways, indicators have in general allowed themselves to be caught up in the eclipsing or replacing of professional networks with marketized information systems (Mirowski 2013). To make this case, we look at two contrasting arenas: the commercialization of Open Access publishing (Muellerleile), and an emerging humanities discipline which lacks commercial potential (Newfield).

In fact, they do not. Metrics have become indissociable from “audit culture” (Shore and Wright, 2015), and the critique of audit culture is a mature field with at least two decades of work behind it. Shore and Wright (2015, 430-31) offer a summary description of its effects:

1. Loss of organizational trust (O’Neill 2002; Power 1994);
2. Elaborate and wasteful gaming strategies (House of Commons 2004; Shore and Wright 2000; Wright 2009);
3. A culture of compliance and large compliance costs, including the appointment of new specialists preoccupied with creating positive (mis)representations of performance (Miller 2001);
4. Defensive strategies and blamism that stifle innovation and focus on short-term objectives over long-term needs (Hood 2002);
5. Deprofessionalization, a disconnect between motivation and incentives, lower employee morale, and increased stress and anxiety (Bovbjerg 2011; Brenneis, Shore, and Wright 2005; Wright 2014);
6. “Tunnel vision” and performing to the measure, with a focus solely on what is counted, to the exclusion of anything else (Townley and Doyle 2007);
7. And the undermining of welfare and educational activities that cannot be easily measured (King and Moutsou 2010).

Since the authors of guidelines for the use of metrics (e.g. Hicks et al) would likely object to any of these seven features, not to mention their combination, why has audit practice not only survived but thrived despite this critique?

Our paper offers two cases in which we explore the marketization imperative. The systems of measuring and ranking have become objects of economic development themselves, driving innovation and the construction of new firms and markets, many of which cross over the boundaries of what was once a more autonomous university (Komljenovic and Robertson 2016). This is exacerbated by at least two things: austerity or the constant push to “do more with less,” and the related managerial fascination with “big” data and “evidence driven” decision making as offering a whole new level of economization. In other words, while there is a political or ideational project at work in the process of neoliberalizing universities, there is also a material restructuring around the construction of new commodities, new markets for those commodities, and new management structures to control this new economy.

Our first case is Open Access publishing. OA has evolved through three main arguments that continue to shape the discourse today. The first is that properly functioning democratic societies are dependent upon the free circulation of knowledge (Stiglitz 1999). The argument is particularly vehement where public resources fund research, but must also pay to access the results. The second and related argument suggests that academic research and knowledge are major drivers of innovation (Stiglitz 1999, Howells et al. 2012). As such the results should be easily accessible to fuel economic competitiveness and growth. The third was a reaction to the “serials crisis” (House of Commons 2004) or the inability of university libraries to afford access to all the journals deemed necessary, coupled with what are widely seen as unjustified profits accumulated by publishers (Ciancanelli 2007). OA is to rectify access and cost problems and enable the circulation of knowledge that supports both innovation and democracy.

But both the “gold” and the “green” variants of OA have problems. In particular, they have trouble justifying the expense of peer review and the maintenance of a stable knowledge archive. They are encouraging moves toward a “publish first, filter second” mode, and then resolving questions of relevance, evidence, and overall quality posed by the lack of prior review by developing and selling bibliometric tools to filter content after relatively unreviewed publication. These tools serve two broad functions, although they are often co-constitutive. First, publishers are developing technologies that categorize, codify, and measure research and researchers. And second, publishers are using these tools to enclose, and sell meta-data about research. Through an evolution of internet media provision, what Mansell (1999) calls the “scarcity-abundance dialectic,” the largest academic publishers are losing control of content, but at the same time becoming massive data aggregating corporations.

The assumption that is built into most advocacy of open access is that scientific, or just academic knowledge must be free for the public to read, if not free to put to use in any way they see fit. On the surface, this seems quite reasonable. But in order to achieve this within the current technical-economic conjuncture, the existing structures that organize academic knowledge and make it meaningful are being dismantled. In turn this is threatening to further alienate universities from the very people who open access advocates claim to have in their interest. Stated differently, capital in the form of subscription based publishers, have historically enclosed knowledge behind pay walls and copyright, but in the process they also helped to make knowledge robust and meaningful. Open Access advocates might argue that in a world of open knowledge and data, a simple Google search will solve the problem by identifying the most popular, well-connected, or most trusted research. The problem is that the for-profit publishers are ahead of this game. They are working very hard to set the rules by which Google or Mendeley or Scopus will identify the “best” academic research on any given topic. Put another way, the information structures of the Internet are not flat. They are always already filtered, curated, and uneven. Furthermore, the algorithms that control these searches are increasingly hidden from human view, or are too complicated for humans (e.g. academic researchers, academic administrators, the broader public) to understand without the aid of digital technology (Gitelman 2013, Pasquale 2015).

In spite of its potential to make knowledge more accessible, OA is being marketized in a way that will re-trigger the critique of audit culture, which details the negative institutional effects of separating research management from research work. OA can

be imposed on researchers by institutional authorities, who will adapt or even embrace it. These do not reduce the intellectual or institutional costs that the critique chronicles.

Our second case is an emergent interdisciplinary U.S. field called Critical University Studies (CUS). CUS was developed and identified by humanities scholars, particularly from literary and cultural study (LCS) (Newfield 2008; Williams 2012). Its findings are of potential importance for the sociocultural life of global universities, but have no obvious market potential. The study of higher education has been marginal in U.S. education departments and schools, and has been shared among historians, sociologists, anthropologists, philosophers, accountants, management professionals, science and technology scholars, and literary critics, among others. How would an interdisciplinary terrain, rooted in informal social networks, be tracked by bibliometrics as it emerges into a new para-discipline?

This paper will report on the results of a comparison of Thomson Reuters Citation Index results to an analysis of the professional circulation of one of the key concepts developed by CUS, “cross subsidies” for extramural research conducted at U.S. universities. Conventional campus wisdom in the 2000s was that Science, Technology, Engineering and Mathematics (STEM) research generated positive revenues that subsidized money-losing fields in the arts, social sciences, and humanities. The literature on this topic was sparse and often trapped in local contexts, such as university-specific reports (UCPB 2010) or findings for membership audiences (COGR 2013). This situation began to change after 2008 through a small number of non-standard statements and writings—a quotation of a budget expert in the *New York Times* (Lieber 2009) which prompted a letter to a university president and a specialty-newspaper article (Watson 2010), and an article about budgeting in a literary journal (Newfield 2009). Many lectures were given that included references to this information (Newfield delivered 120 lectures from late 2011 through mid-2015). Over a period of 6-7 years, the conventional wisdom on cross-subsidies was reversed in arts and humanities circles, and at least challenged in STEM circles. But much of this change was social and word-of-mouth, contained in discussions and debates in department meetings and at dinner parties, and prone to resurface in publication sans attribution (Dinsman 2016). We posit that CUS represents a common pattern for emerging professional knowledge, in which such knowledge adapts to local contexts and evolves as it migrates through affinity networks, while tending to shy away from the high-profile venues most concerned with their reputation and markets. As a result, emerging knowledge evades marketization, and can be undermeasured. Its circuits function more like an artistic avant-garde or a musical subculture (Hebdige 1979) than like a citational network.

Through empirical descriptions of the circulation of CUS knowledge, we will suggest that core axioms of Garfield bibliometrics do not function normally in a non-marketized system of professional knowledge. Here we do *not* find that citation frequency is a reliable index of cognitive impact (De Bellis 2009); that citation indexes measure the intellectual impact and productivity of individuals and units over time, and thus can be used for quality assurance and other management functions (Hirsch 2005); that the standard concentration of references to a small number of scientists reflects the social actuality of knowledge generation and influence; that influence follows a power law function, or follow Pareto rather than Gaussian distribution (Lotka, 1926; Bradford 1934; Zipf, 1936; De Bellis, 2009). Although the study of literature and the arts has sometimes adopted a belief in the concentration of genius by focusing on canonical

masterworks, this has decreasingly been the case in recent decades, as the humanities fields have studied the circulation of aesthetic and cultural intelligence throughout entire social systems, particularly across cultural differences and around the alleged peripheries.

Our comparison of these two arenas, OA and CUS, will allow us to conclude with suggestions for how bibliometrics might track distributions of knowledge that are neither Gaussian nor Paretian, but informal, subterranean, and democratic. This may be an opportunity for quantitative assessment to part company with audit culture.

REFERENCES

- Archambault, Éric and Étienne Vignola Gagné. 2004. *The Use of Bibliometrics in the Social Sciences and Humanities*. Montréal: Social Sciences and Humanities Research Council of Canada.
- Bovbjerg, Kirsten-Marie, ed. 2011. "Motivation og mismod: effektivisering og stress på offentlige arbejdspladser." Aarhus, Denmark: Aarhus University Press.
- Bradford, Samuel C. 1934. "Sources of Information on Specific Subjects," *Engineering* 137(3550): 85-86.
- Brenneis, Don, Cris Shore, and Susan Wright. 2005. "Getting the measure of academia: universities and the politics of accountability." *Anthropology in Action* 12(1):1-10.
- Ciancanelli, P. 2007. "(Re)producing universities: Knowledge dissemination, market power and the global knowledge commons." In Epstein, D., Boden, R., Deem, R., Rizvi, F., and Wright, S. (eds.) *World Yearbook of Education 2008*, New York: Routledge
- COGR (Council on Governmental Relations). 2014. "Finances of Research Universities." May. http://cogr.edu/COGR/files/ccLibraryFiles/Filename/000000000267/Finances%20of%20Research%20Universities_June%202014.pdf
- Cronin and Cassidy R. Sugimoto, eds. 2014. *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact*. Cambridge, Mass.: MIT Press.
- De Bellis, Nicola. 2009. *Bibliometrics and Citation Analysis: From the Science Citation Index to Cybermetrics*. Lanham, Md.: The Scarecrow Press.
- Dinsman, Melissa. 2016. "The Digital in the Humanities: An Interview with Alexander..." *The Los Angeles Review of Books*, March 27. <https://lareviewofbooks.org/interview/the-digital-in-the-humanities-an-interview-with-alexander-galloway/>.
- Gill, John. 2015. "Taking Appropriate Measures," *Times Higher Education*. 9 July 2015: 5.
- Gingras, Yves. 2014. "Criteria for Evaluating Indicators." In Cronin and Sugimoto, eds. 109-25.

- Gitelman, L. ed. 2013. *“Raw Data” is an Oxymoron*, Cambridge, MA: MIT Press
- Hebdige, Dick. 1979. *Subcultures: The Meaning of Style*. London: Routledge.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). [The Leiden Manifesto for Research Metrics](#). *Nature*, 520, 429–431.
- Hood, C. 2002. “The risk game and the blame game.” *Government and Opposition* 37(1):15–37.
- House of Commons. 2004. “Scientific Publications: Free for all?” House of Commons Science and Technology Committee, Tenth Report of Session 2003-2004. London: Her Majesty’s Stationary Office:<http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/399/399.pdf>
- Howells, J., Ramlogan, R., and Cheng, S.L. 2012. “Innovation and university collaboration: paradox and complexity within the knowledge economy.” *Cambridge Journal of Economics*, vol. 36: 703-721.
- King, Lucy, and Christina Moutsou, eds. 2010. *Rethinking audit cultures*. London: PCCS Books.
- Komljenovic, Janja & Susan L. Robertson (2016): The dynamics of ‘market-making’ in higher education, *Journal of Education Policy*
- Lotka, Alfred J. 1926. "Statistics: The Frequency Distribution of Scientific Productivity." *Journal of the Washington Academy of Sciences* 16 (12): 317-25.
- Lieber, Ron. 2009. “Why College Costs So Much.” *The New York Times*, September 4. <http://www.nytimes.com/2009/09/05/your-money/paying-for-college/05money.html>.
- Miller, Peter. 2001. “Governing by numbers: why calculative practices matter.” *Social research* 68(2):379–396.
- Mansell, R. 1999. “New media competition and access: the scarcity-abundance dialectic”, *New Media and Society*, vol. 1, iss. 2: 155-182.
- Mirowski, Philip. 2013. *Never Let a Serious Crisis Go to Waste: How Neoliberalism Survived the Financial Meltdown*. London ; New York: Verso.
- Newfield, Christopher. 2008. *Unmaking the Public University: The Forty-Year Assault on the Middle Class*. Reprint. Harvard University Press.
- Newfield, Christopher. 2009. “Ending the Budget Wars: Funding the Humanities during a Crisis in Higher Education.” Modern Languages Association: *Profession*.
- NICE. 2013. *Guide to the methods of technology appraisal*. London: National Institute of Health and Clinical Excellence <http://publications.nice.org.uk/pmg9>

- O'Neill, Onora. 2002. "A question of trust: the BBC Reith lectures of 2002." Cambridge: Cambridge University Press.
- Power, Michael. 1994. *The audit explosion*. London: Demos. <http://www.demos.co.uk/files/theauditexplosion.pdf>
- Power, Michael. 1997. 1997. *The audit society: rituals of verification*. Oxford: Oxford University Press.
- Shore, Cris and Susan Wright. 2015. "Audit Culture Revisited: Rankings, Ratings, and the Reassembling of Society." *Current Anthropology* 56: 3, June.
- Stiglitz, J. 1999. "Knowledge as a Public Good." In Kaul, I, Gruenberg, I. and Stern, M.A. *Global Public Goods: International Cooperation in the 21st Century*. Oxford: Oxford University Press
- Townley, Barbara, and Rosemary Doyle. 2007. "Performance measurement." In *Blackwell Encyclopedia of Sociology*. George Ritzer, ed. Blackwell Reference Online. http://www.blackwellreference.com/subscriber/tocnode.html?idpg9781405124331_chunk_g978140512433122_ss1-17
- UCPB (University Committee on Planning and Budget). 2010. "Choices Report." Oakland: University of California Office of the President. <http://senate.universityofcalifornia.edu/reports/ucpb.choices.pdf>
- Watson, Robert N. 2010. "The Humanities Really Do Produce a Profit." *The Chronicle of Higher Education*, March 21. <http://chronicle.com/article/The-Humanities-Really-Do/64740/>.
- Williams, Jeffrey J. 2012. "Deconstructing Academe." *The Chronicle of Higher Education*, February 19. <http://chronicle.com/article/An-Emerging-Field-Deconstructs/130791/>.
- Wright, Susan. 2014. "Knowledge that counts: points systems and the governance of Danish universities." In *Under new public management: institutional ethnographies of changing front-line work*. Dorothy Smith and Alison Griffith, eds. Toronto: University of Toronto Press.
- Wright, Susan. 2005. "Processes of social transformation: an anthropology of English higher education policy." In *Pædagogisk Antropologi—et Fag I Tilblivelse*. J. Krejsler, N. Kryger, and J. Milner, eds. Pp. 185–218. København: Danmarks Pædagogiske Universitets Forlag.
- Wouters, Paul. 2014. "The Citation: From Culture to Infrastructure." In Cronin and Sugimoto, eds. 47-66
- Zipf, George K. 1935. *The Psycho-Biology of Language: An Introduction to Dynamic Philology*. London: Routledge, 1936.



On the Peripheries of Scholarly Infrastructure: A look at the Journals Using Open Journal Systems¹

Juan Pablo Alperin^{*}, Kevin Stranack^{**} and Alex Garnett^{**}

*juan@alperin.ca

Canadian Institute for Studies in Publishing, Simon Fraser University, 515 West Hastings Street, Vancouver, 6B 5K3 (Canada)

Public Knowledge Project, Simon Fraser University, 8888 University Drive, Burnaby, V5A 1S6 (Canada)

**kstranac@sfu.ca; garnett@sfu.ca

Public Knowledge Project, Simon Fraser University, 8888 University Drive, Burnaby, V5A 1S6 (Canada)

ABSTRACT

Although there have been calls for scholarly infrastructure to be inclusive, new layers of infrastructure are built without a clear understanding of the breadth of scholarly journals that lie on the peripheries of the existing infrastructure. In the hopes that future infrastructure can take a wider range of journals into account, this paper presents the results of an effort to track the number, location, and rate of publication of journals using Open Journal Systems, an open source manuscript management and publication system built by the Public Knowledge Project. The method employed, which involves a combination of scanning weblogs, scraping webpages, and harvesting metadata, has yielded an estimated 9,828 journals that have collectively published 2,565,300 articles since 1990. These journals are distributed across 136 countries on 6 continents, and, in 2015, around a fifth of the OJS journals were published in low or low-middle income countries, and over a third in upper-middle income countries, suggesting that the majority of OJS journals are on the “periphery” of today’s global scholarly infrastructure. As infrastructure and services continue to be developed, this paper argues, it is necessary to look to such journal so that the infrastructure that is built can be developed in a way that is truly inclusive of the global nature of scholarship.

BACKGROUND

The Public Knowledge Project (PKP) is a research and development initiative of Simon Fraser University and Stanford University, with a focus on understanding and building enhanced modes of scholarly communication that facilitate open access, high quality publishing, and local capacity building and participation.

One of the most significant contributions from PKP has been the development of its free, open source Open Journal Systems (OJS) software.¹ OJS is a professional journal publishing platform that is easy to download, install, and operate with minimal server requirements. OJS allows for online submissions, peer review, copyediting and layout, and publishing. It also provides connections to indexing (e.g., PubMed, DOAJ), digital identifiers (e.g., CrossRef, ORCID), and preservation services (e.g., LOCKSS, Portico).

Because OJS is free and is designed to maximize efficiencies for publishing activities, publishers have been able to launch journals that would otherwise have been financially untenable. Although some existing publishers have made use of the system, the majority of

¹In the interest of full disclosure, it should be noted that the authors of this work are all affiliated with the Public Knowledge Project, the creators of Open Journal Systems.

OJS users are new to publishing and are based in academia, either through library publishing programs or by individual "scholar-publishers," determined to build communities of interest around their research areas (Edgar & Willinsky, 2010). Most are small-scale, often operating with in-kind contributions and minimal budgets (Edgar & Willinsky, 2010). This type of use suggests an increased participation in scholarship, both in terms of who publishes and who read.

In a sense, OJS can already be said to form part of the scholarly infrastructure. It provides a layer over which scholarly activity takes place, and it brings the use of standards and best practices to those activities. It acts as one layer of infrastructure especially for those journals who would be not formally part of the scholarly infrastructure

METHODS

We have sought to use automated methods for identifying OJS journals, and subsequently use a combination of web scrapingⁱⁱ and the journal's OAI PMH endpointⁱⁱⁱ to collect publicly available information about the journal, including article metadata. Because this process is entirely automated, it can be continuously run, yielding a dataset of journals and articles data that will eventually include longitudinal data. These data are then processed to identify the number of journals, articles, and their geographic location.

In a latter phase, yet to be completed (and thus not presented here) these data will be complemented with an online survey, loosely based on a similar study of OJS journals done by Edgar & Willinsky (2010), and targeted at the email addresses collected through the first phase data collection.

The first challenge in studying these journals is to learn of their existence. Because the OJS software is open source, each journal or publisher can install the software on their own server, without ever needing to register with PKP or elsewhere. Only once a journal's web address (URL) has been identified is it possible to learn how many articles it publishes in any given year, its geographic location, and its identifying information and other metadata. To collect this data, we have devised a necessarily complex method of extracting and processing the information (Figure 1).

Figure 1. Steps to collect and process OJS journal information

1. Process the PKP website logs for referrer URLs that 'look like' OJS journals
2. Attempt to contact the OAI PMH URL corresponding to the journal URL (following known OJS URL patterns) to verify if it is an OJS journal
3. Save the repository identifier, I.P. address, OJS version number
4. Identify all the journals for this installation using the OAI verb "ListSets"
5. Save the journal name, and journal contact email address from the OAI response for later use
6. Add known OJS OAI URLs to an instance of the PKP Harvester
7. Look up the journal's country
8. Collect the article metadata for every journal using OAI PMH
9. Process the article data to identify number of articles published per year, the country of origin of the journal, etc.

Note: Code for step 1 can be found at: <https://github.com/pkp/ojsstats/blob/master/pkp-log-parser.php>; Code for steps 2-5 can be found at: <https://github.com/pkp/ojsstats/blob/master/checkOJS.py>; Code for step 6 can be

found at: <https://github.com/pkp/ojsstats/blob/master/harvesterInsert.py>; Code for step 7 can be found at: <https://github.com/pkp/ojsstats/blob/master/countryLookup.py>; Step 8 is handled by the PKP Harvester, which can be found at: <https://github.com/pkp/harvester>; Code for step 9 can be found at: <https://github.com/pkp/ojsstats/blob/master/statscrunch.py>. The process is coordinated by a “cron” job, which can be found at: <https://github.com/pkp/ojsstats/blob/master/cron.sh>. Detailed description of the each step can be found in the README: <http://github.com/pkp/ojsstats>.

There are several known ways in which the above system can fail to correctly identify and query a valid OJS journal. It is possible, for example, that a journal removes the links to the PKP website from their installation altogether, rendering it invisible to our logs; similarly, it is possible that journals have modified the code to change the way URLs are structured, making them difficult to correctly identify; it is also possible that journals have disabled their OAI endpoint, thereby invalidating the way in which we verify an OJS installation and subsequently collect data; or, a journal may simply be unavailable online at the time when we attempt to crawl it, leading our system to believe that the site no longer exists. Any of these circumstances, and possibly others, would result in us undercounting the number of OJS journals.

While we realize there are simpler ways of collecting this data about OJS journals (i.e., the data could be pushed from the OJS installation to a centralized system), it has always been PKP’s approach to give control of the data to the journals, and to ask for nothing in return for using the software. In recent years, however, PKP has recognized the need to better understand its journal community, and, as a result, has included a “phone home” feature that provides PKP with a minimal set of publicly available data (with an opt-out option). This feature has only been available since 2015, so it was still necessary to develop the methods outlined above to produce the results that follow.

RESULTS

We identified 6,271 installations of OJS with some content, spread across 136 countries on 6 continents. These installations collectively host 9,828 “journals” that meet our arbitrary threshold of at least 18 articles published in the previous two years.^{iv} At the time of writing, there were 7,491 journals that met this threshold for 2015, and 9,315 that met it for 2014 (some journals appear to add content with a delay in publication resulting in the appearance of a drop-off in numbers in 2015). Of these, the top 3 countries, by number of journals published using OJS with recent content, are 1,426 in Brazil, 1,075 in Indonesia, and 912 in the US (Figure 2). No other country had more than 500 journals that met the criteria in 2015. Interestingly, Latin America (led by Brazil) publishes approximately one third of all OJS journals.

Naturally, the corresponding country income level accompanies the geographic distribution. Approximately 42% of the journals are published by high-income countries, 38% by upper-middle income countries, and the final 20% by low or low-middle income countries (Table 2).^v

These journals have published 2,565,300 articles since 1990. As might be expected, more articles are published using the software in more recent years, with over 300,000 articles published in 2013 and 2014, and over 250,000 in 2012 and 2015 (Figure 3). This is due primarily to an increase in the number of journals that have gone (and stayed) online over time (Figure 4). Given that the software was originally released in 2001, it seems that journals come online, add some amount of back-content (archives), and then continue to publish. In

recent years (since 2010), OJS journals that meet the 18-article threshold have published 29.8 articles (up from 25.6 in the 15 years prior) (Figure 5).

Figure 2. Location of OJS journals in 2015

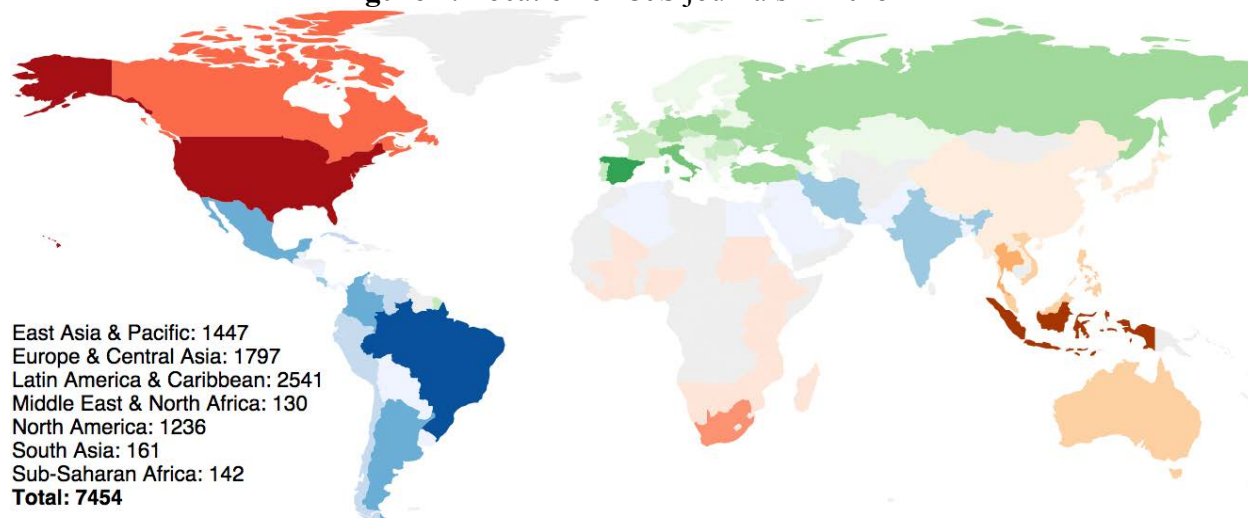


Table 2. Country Income-Level of OJS Journals for 2015

Income Category	Number and percent of journals (2015)	Number and percent of articles (2015)	Average Number of Articles per journal (2015)
Low	53 (0.7%)	1,225 (0.4%)	23.1
Lower-middle income	1,525 (19.7%)	68,571 (23.8%)	46.6
Upper middle income	2,832 (37.8%)	94,945 (33.0%)	33.5
High income	3,129 (41.8%)	123,211 (42.8%)	39.4

Figure 3. Number of Articles Published in OJS Journals By Year

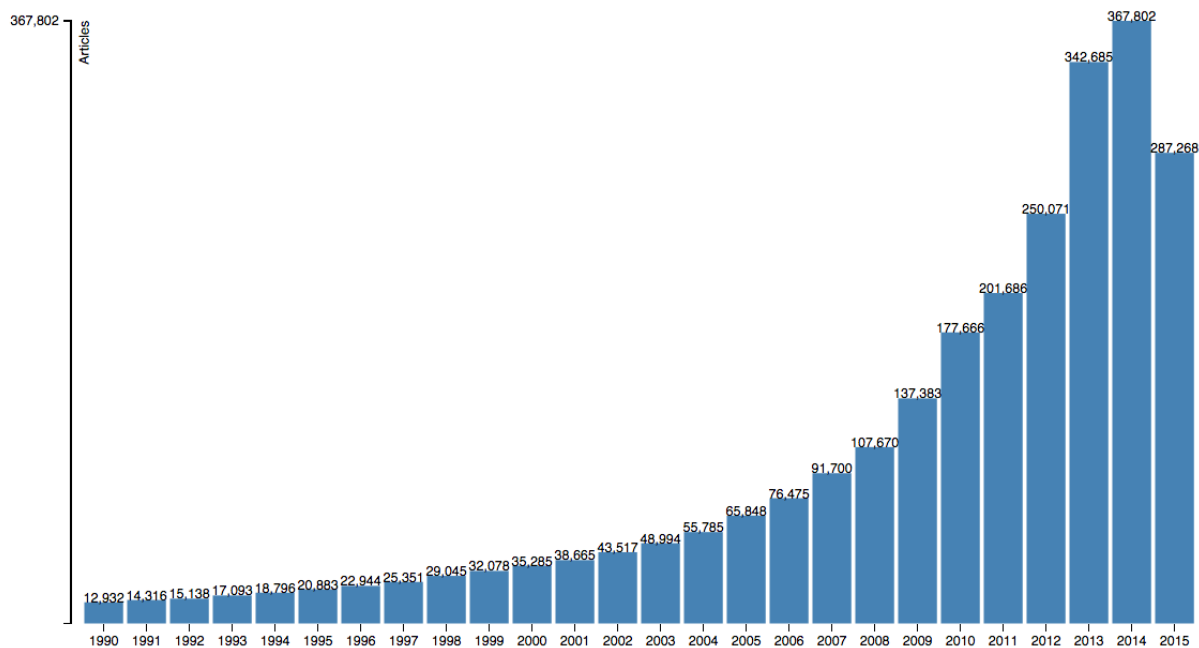


Figure 4. Number of Journals Using OJS by Year

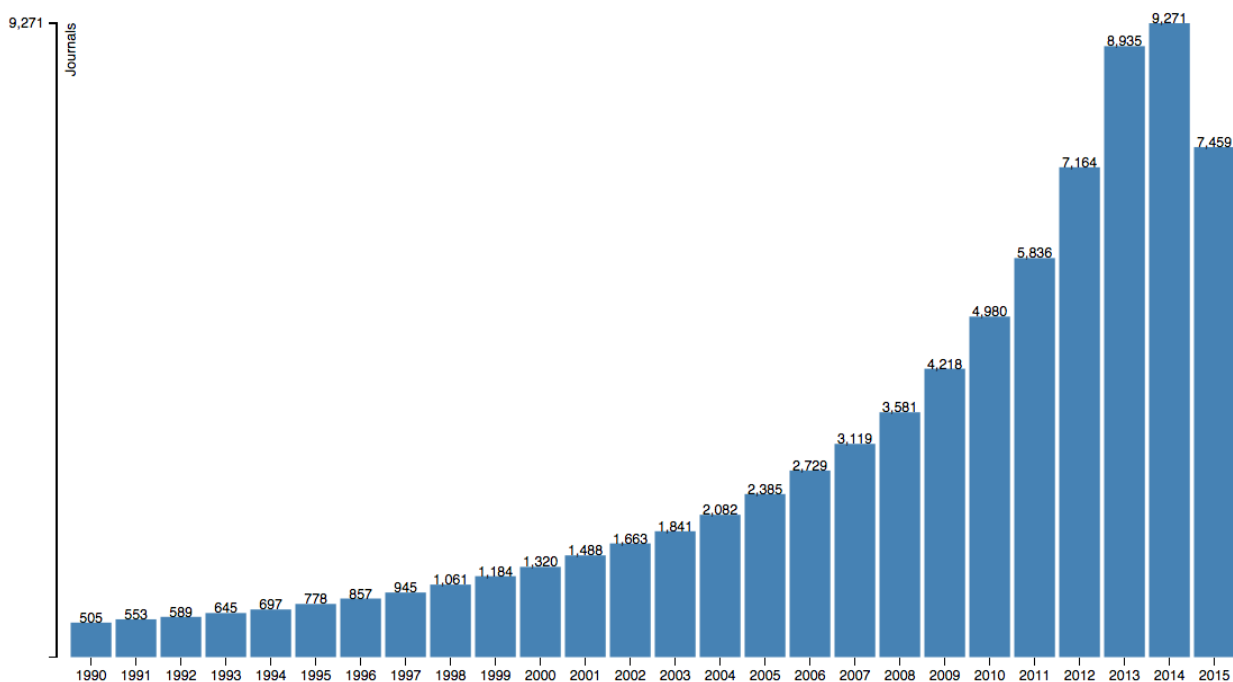
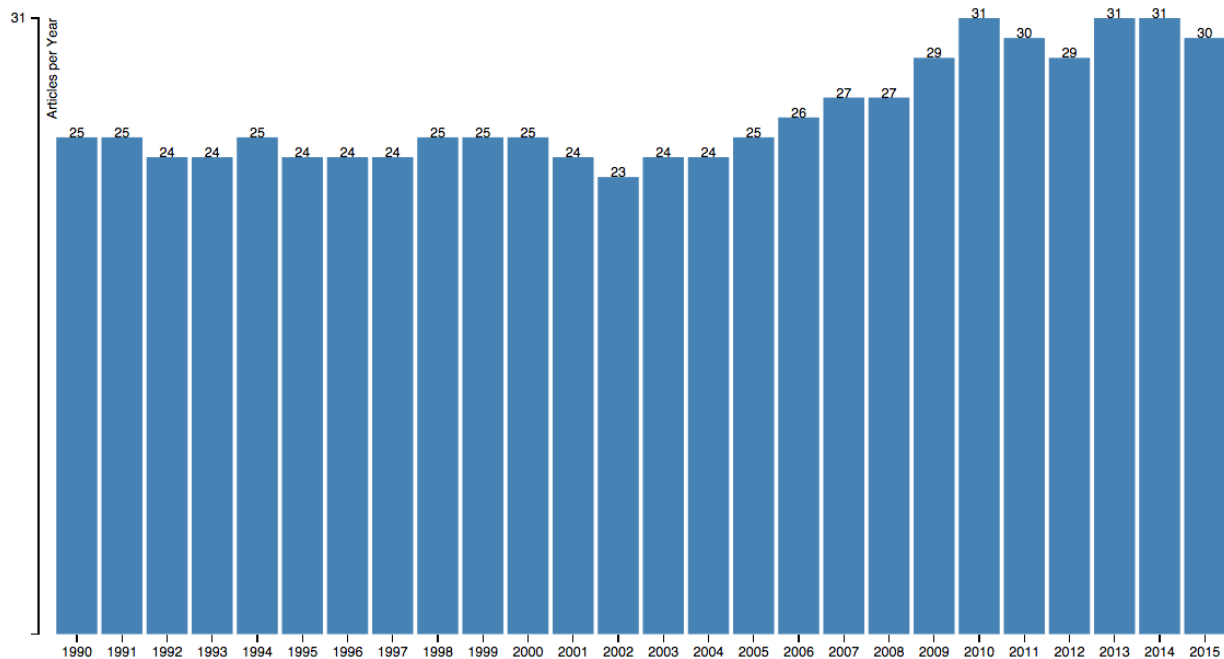


Figure 5. Average Number of Articles Published per Year in OJS Journals

DISCUSSION AND FUTURE WORK

Because of the imperfect criteria and roundabout method for collecting the data, these numbers should be treated as estimates. However, although we are surely counting some instances that are not really journals as have been traditionally understood, we are also certain that we have missed some instances altogether, and are reasonably comfortable that they provide a sufficiently accurate estimate to provide a better understanding of the type of journals that are in need of access to the scholarly infrastructure. Future studies and improved methods, including the better automated data collection, will allow us to assess these estimates.

The geographic extent and sheer number of journals and articles—most of which are not likely to be found in international citation databases—should serve to open our eyes to the need to think more broadly about who is actively trying to communicate scholarship.

The extent to which these journals are currently outside the scholarly infrastructure cannot be overstated. Not only are these journals excluded from major citation databases (there is no need to test their inclusion, as the sheer volume of journals from outside North America and Europe dwarf the total number of journals from these regions in Web of Science and Scopus), they are also lacking some of the basic elements of the scholarly infrastructure, such as DOIs. As of November 2015, there were only 837 CrossRef members that used OJS (CrossRef, personal communication). These members collectively minted 701,622 DOIs (out of over 2.5M articles).^{vi} Even under the generous assumption that each DOI minted corresponds to an article (i.e., not supplementary files, data, etc.), this amounts to less than a third of the total articles published with OJS journals.

The geographic and economic distribution of access to elements of the scholarly infrastructure, such as DOIs, remains to be studied.^{vii} It is nonetheless striking (although likely coincidental) that the number of DOIs minted by CrossRef members using OJS in 2015

(127,026) coincides with the number of articles published by OJS journals from high-income countries that year (123,211).^{viii} While PKP works towards improving the adoption of DOIs for these journals and CrossRef works to revamp its small-publisher tools, the reality is such that the scholarly metrics and indicators that can be calculated today do not consider these journals on the peripheries of the scholarly communication infrastructure.

As can be seen from the complexity of the data collection described above, deriving metrics from these small, decentralized publishers is a challenge. Attempting to extend our efforts beyond the journals published with OJS would make the task nearly impossible. In this way, OJS is itself a piece of the infrastructure, one that allows us to account for over 2.5M articles and gives us the opportunity to learn more about them.

If these journals were part of the larger metrics infrastructure and systems, it would fundamentally change the peripheral nature of the scholars and scholarly work contained therein. As these figures highlight, there are literally millions of articles published from this periphery. The over 20% of OJS journals in low and low-middle income countries (corresponding to nearly 25% of all OJS-published articles) need to be included if the indicators and metrics are to be truly representative of the scholarship that is out there.

Being inclusive in the indicators and metrics infrastructure is necessary if we wish to use these indicators to understand how scholarship is carried out around the world. Being inclusive, however, is not as simple as putting out an open invitation to participate. It is also necessary to lower the barriers to access by making sure the infrastructure is suitable and adapted to the needs of everyone. To do that effectively, we need to know who we are trying to include, so that they can be consulted and can participate from the beginning as the infrastructure and metrics are designed and built. The research described here is a first look at many of the thousands of journals who are currently not being included.

This work is very much ongoing. In the coming months, we will conduct a survey of these journals to learn about the practices, challenges, funding, and other information that will be directly relevant to those seeking to build truly inclusive scholarly infrastructure. While this first phrase has given a general sense of the scale and extent of journals, it is only the beginning of a longer process of trying to understand the global research landscape.

REFERENCES

Alperin, J. P. (2013, October 11). Are ALMs/altmetrics propagating global inequality? Retrieved from <https://speakerdeck.com/jalperin/altmetrics-propagating-global-inequality>

Alperin, J.P. (2014a). Altmetrics could enable scholarship from developing countries to receive due recognition. *LSE Impact Blog*. Retrieved from: <http://blogs.lse.ac.uk/impactofsocialsciences/2014/03/10/altmetrics-for-developing-regions/>

Alperin, J. P. (2014b). South America: Citation databases omit local journals. *Nature*, 511(7508), 155–155. <http://doi.org/10.1038/5111155c>

Bilder, G., Lin, J., & Neylon, C. (2015). Principles for Open Scholarly Infrastructures-v1. <http://doi.org/10.6084/m9.figshare.1314859.v1>

Cetto, A. M., & Alonso-Gamboa, J. O. (1998). Scientific periodicals in Latin America and the Caribbean: a global perspective. *Interciencia*, 23(525), 84–93. Retrieved from http://www.interciencia.org/v23_02/cetto.pdf

Collazo-Reyes, F., Luna-Morales, M. E., Russell, J. M., & Pérez-Angón, M. (2008). Publication and citation patterns of Latin American & Caribbean journals in the SCI and SSCI from 1995 to 2004. *Scientometrics*, 75(1), 145–161. <http://doi.org/10.1007/s11192-007-1841-6>

Edgar, B. D., & Willinsky, J. (2010). A Survey of the Scholarly Journals Using Open Journal Systems. *Scholarly and Research Communication*, 1(2), 1–22.

Haustein, S., Sugimoto, C., & Larivière, V. (2015). Guest editorial: social media in scholarly communication. *Aslib Journal of Information Management*, 67(3). <http://doi.org/10.1108/AJIM-03-2015-0047>

Russell, J. M., Madera Jaramillo, M. J., & Shirley, A. (2009). El análisis de redes en el estudio de la colaboración científica. *Redes: Revista Hispana Para El Análisis de Redes Sociales*, 17(2), pp. 39–47. Retrieved from <http://dialnet.unirioja.es/servlet/articulo?codigo=3165613&orden=243060&info=link>

Terra Figari, L. I. (2008). Diseminación del Conocimiento Académico en América Latina. In *Antropología Social Y Cultural En Uruguay 2007*. Montevideo, Uruguay: UNESCO. Retrieved from http://www.unesco.org.uy/shs/fileadmin/templates/shs/archivos/anuario2007/articulo_15.pdf

ⁱ <http://pkp.sfu.ca/ojs>

ⁱⁱ Using a script to make web requests, like a browser would, to access content from the journal pages.

ⁱⁱⁱ OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) is a standard used by many journal and repository systems to expose metadata in a machine-readable format.

^{iv} This 18 articles in the previous two years is based on Canada’s Social Science Research Council’s Aid to Scholarly Journals’ guideline that stipulates 12 peer reviewed articles in the previous two years (see http://www.sshrc-crsh.gc.ca/funding-financement/programs-programmes/scholarly_journals-revues_savantes-eng.aspx#a5 for details). We assume, therefore, that two thirds of articles are peer reviewed to arrive at an 18 article threshold). This minimum was set in an effort to filter out test installations, or installations that have just gone online and have not fully gotten underway. Since we are not verifying any other criteria, the word “journal” is therefore used loosely.

^v The income levels are drawn from the World Bank classifications, with both high-income categories (OECD and non-OECD combined) and the low and lower-middle income categories combined.

^{vi} Note that a member could correspond to more than one journal.

^{vii} Ironically, the most suitable way of cross-checking the OJS database and others would be to use DOIs.

^{viii} Even if this were not a coincidence, and the majority DOIs were minted by journals from high-income countries, it should be noted that DOIs minted in 2015 do not necessarily correspond to content published in 2015, so it is unlikely there would be a one-to-one correspondence.



Why researchers publish in journals not indexed in mainstream databases: training, bridging and gap-filling¹

Diego Chavarro^{*}, Puay Tang^{*} and Ismael Rafols^{*,**}

^{*} diego.chavarro@sussex.ac.uk, p.tang@sussex.ac.uk

SPRU (Science Policy Research Unit), University of Sussex, Brighton, BN1 9SL, England

^{**} i.rafols@ingenio.upv.es

Ingenio (CSIC-UPV), Universitat Politècnica de València, València, 46022, Spain
Observatoire des Sciences et Techniques (OST-HCERES), Paris, France

ABSTRACT

Although journals indexed in mainstream Journal Indexing Systems (JIS), i.e. Web of Science (WoS) and Scopus, have more visibility, reputation and are more valued in evaluation, in developing countries researchers keep publishing in journals not indexed in mainstream databases, but indexed instead in alternative JIS such as Scielo or RedALyC. The conventional explanation to this behaviour is that developing countries' research often does not have sufficient quality so as to be published in journals indexed by WoS or Scopus. We conducted 30 interviews to researchers in Colombia working in chemistry, agricultural sciences, and business and management asking the reasons for publishing in different types of journals, in particular those indexed by mainstream and alternative JIS. The answers provided by the researchers were classified into three reasons. The first reason is that journals in alternative JIS offer a space for *training* in publishing, both as an introduction to academic publishing and as a step-stone towards publishing in WoS. The second reason is that journals indexed by Scielo and RedALyC have a *knowledge-bridging* function, providing a link between articles covered by mainstream JIS and articles of regional communities with limited access to WoS or Scopus journals. The third reason is that alternative JIS journals have a *knowledge-gap filling* function, allowing the publication of topics that are not well covered in WoS-indexed journals, such as locally relevant agricultural products or regional history. We conclude that scientometric indicators based on mainstream JIS underrepresent the contribution of research from developing countries – as they do not value these training, knowledge-bridging and gap-filling functions. We discuss these findings in the light on universalistic versus particularistic conceptualisations of science.

INTRODUCTION

Why do researchers in developing countries such as Colombia keep publishing in journals not indexed by WoS or Scopus, given their low visibility, low reputation and that they are lowly valued in evaluation exercises? The explanation implicitly held by many evaluation or scientometric experts is that if they could, all researchers would publish in journals indexed in WoS, which are those with the highest quality. This belief follows from the ideas on research quality of Eugene Garfield, founder of WoS. He argued that “the significant scientific literature appears in a small core of journals” (Garfield, 1996). According to him, this core

¹ This work is based on insights from one chapter of Diego Chavarro's thesis in SPRU (University of Sussex, submitted in 2016) supported by a fellowship from Colciencias, Colombia. PT and IR provided advice as his supervisors.

was composed of around 150 journals that “account for half of what is cited and a quarter of what is published in WoS” (Garfield, 1996). The concept of core journals has been used to determine and justify the coverage of WoS. Basically, the aim of WoS is to select a portion of scientific journals characterised by their high scientific impact and their compliance with editorial standards. These characteristics can be considered as universalistic: they could be achieved by any journal regardless of its language, discipline, or country of publication. Within this rationale, the journal coverage of WoS is considered objective and the exclusion of journals is seen as justified.

If this is the case, why should researchers publish outside WoS or Scopus? The development of alternative Journal Indexing Systems (JIS) such as Scielo or RedALyC suggests that non-mainstream journals fulfil functions that are valued by researchers and policy-makers in regional contexts such as the Ibero-American. An analysis of WoS’ and Scopus’ coverage shows that their coverage is particularistic, meaning that geographic, linguistic, and disciplinary biases have an important impact on journal selection decisions (Chavarro, 2016, unpublished). As a result one can expect that publication in journals indexed in alternative JIS have various valuable functions beyond or rather than publishing “low quality” research (i.e. research that is not perceived as making a significant contribution to knowledge by peers in ‘global’ scientific communities).

What are the functions performed in the scientific system of a developing country provided by publications in journals indexed by alternative JIS? In this article we investigate these functions by examining the reasons reported in thirty interviews by researchers in Colombia for publishing in diverse journals.

Colombia as a case study

We use Colombia to examine the publishing practice of researchers in developing countries with a growing number of publications in mainstream and alternative JIS. Colombia is classified as an upper-middle income country by the OECD and usually also classified as an S&T developing country (Ordóñez-Matamoros et al, 2010). It is an important producer of scholarly journals in Ibero-America. It can be compared on its production of journals to Brazil, Mexico, Chile, Peru, Venezuela, and Cuba. As in other Ibero-American countries, most of its scholarly publishing houses are higher education institutions. However, few journals produced in Colombia are covered by WoS.

Additionally, a good number of these journals are from the social sciences and are published in Spanish. This means that Colombia has multiple disadvantages in terms of coverage by WoS. At the same time, scientists working for Colombian organisations have increased their production in journals indexed by WoS, which is a trend in Ibero-America (Lemarchand, 2012). This shows two phenomena happening in parallel: the first is the increasing production of journals indexed by alternative JIS; the second is the growing number of papers in WoS-indexed journals by Ibero-American researchers. Ibero-American researchers create these phenomena by their decisions on where to publish – WoS, alternative JIS or both. This makes them an essential source of information on why alternative JIS develop. Their position as researchers in a peripheral country to WoS makes this case valuable to understand the development of alternative JIS.

METHODOLOGY

Data source and sample

30 interviews were conducted from May to September 2013 in Colombia. The sample of researchers was taken from three different disciplines, namely chemistry, agricultural sciences, and business and management. The main reasons to choose these disciplines are extent of coverage and context of application. In terms of coverage, chemistry is generally well covered by WoS while the other two are not. This could imply a lesser need for alternative JIS in chemistry as compared to the other two disciplines.

The researchers in the sample have a variety of backgrounds that are shown on table 1:

Table 1. Distribution of researchers interviewed

Sector	Private university	19
	Public university	11
Experience	Senior	17
	Junior	13
Gender	Women	9
	Men	21
Nationality	Colombian	26
	Other	4

Importantly, these researchers exhibit different publication patterns in journals covered by WoS, Scopus, Scielo, and RedALyC. They were identified using CvLAC. This is a Curriculum Vitae database managed by Colciencias, the main public funding agency for science in Colombia. The criteria to select researchers were based on (1) those participating in a research group endorsed by a Colombian organisation certified by Colciencias², (2) having a PhD, and (3) having an individual production of at least three papers in the last ten years. In actuality most of the interviewees have five or more papers. We contacted 60 researchers in total, and conducted 30 formal interviews with them - ten for each discipline.

Interview protocol

The interview program was intended to answer the research question: *why do researchers publish in journals indexed by alternative JIS?* The interviews followed a semi-structured, open-ended questionnaire. A final questionnaire that we grouped into five main topics:

1. Reasons to publish research.
2. Explanation of the publication patterns of researchers in terms of JIS.
3. Use of Scielo, RedALyC, WoS, and Scopus in research.
4. The “value” of Scielo, RedALyC, WoS, and Scopus for their publications.
5. The future of JIS, recommendations, comments.

Twenty-eight of the interviews were recorded. We used the method known as thematic analysis, which consists of taking notes while interviewing, then journalizing the notes as soon as the interview is finished, listening to the audio files, identifying categories, and validating the categories found through a second review (Braun and Clarke, 2006). The

² In Colombia, in order to be recognised as a research group by Colciencias, the supporting organisation has to confirm it formally. This is known as endorsement.

responses were complemented by secondary data sources. Specifically, the CVs of the researchers in the sample, the examination of Scielo, RedALyC, WoS, and Scopus, and the analysis of specific papers mentioned by them.

RESULTS

The analysis of the interviews suggests three sets of reasons for publishing in journals not indexed in the mainstream databases. Details with quotes of the reasons will be provided in full article, but for lack of space, here we provide the distillation of the insights in terms of three motivations or reasons. We refer to them as *training*, *knowledge-gap filling*, and *knowledge-bridging*.

Training

Journals in alternative JIS-indexed are perceived as “transit stations” towards WoS-indexed journals or training arenas for initiation into publication in WoS-indexed journals. There are two bases for this argument.

- a) **Journals in alternative JIS are used as training for researchers to publish in WoS-indexed journals:** The experience gained by publishing in alternative JIS-indexed journals increases the skills of researchers to publish in WoS-indexed journals. The papers they publish in alternative JIS-indexed journals incorporate this feedback, which contributes to improving the robustness of other papers that will be submitted to WoS-indexed journals in English.
- b) **These journals are also used to introduce PhD students to academic publishing in their own language:** researchers encourage PhD students to look for literature and to publish papers in alternative JIS-indexed journals as part of their formation as academics. This is different from item (a) above in that the ultimate aim is not to publish in WoS-indexed journals, but to initiate new researchers into publishing. Doctoral students also get acquainted with the peer review system, regardless of their success in publishing or their future publication patterns.

In this sense, Scielo and RedALyC are seen as a means, whereas WoS is seen as the goal. This perception comes from the idea that there is a sequential publishing pattern in a researcher’s career: from non-indexed through alternative JIS to mainstream JIS-indexed journals. This is a universalistic understanding of stratification of research quality. Consequently, from this perspective alternative JIS appear to be less important than mainstream JIS. For this reason, some researchers send their “best” contributions to WoS or Scopus and their “second best” papers to alternative JIS-indexed journals because they see less value in the latter. These papers add to the number of documents covered by alternative JIS, contributing to their growth, but are perceived as having less worth than those published in mainstream JIS.

Knowledge-bridging

By knowledge-bridging we mean that publishing in alternative JIS provides a link between articles covered by mainstream JIS and “local” communities with limited or no access to it. Mainstream JIS articles are published in journals based in the UK, the USA or the Netherlands, written in English, and generally require payment for access. From the examples in the interviews (such as the use of business and management papers in the classroom, or the

linguistic differences between Scielo and WoS in passion fruit publications), we made a list of specific mechanisms through which knowledge-bridging is achieved:

- c) **Knowledge adaptation:** The adaptation of knowledge happens when certain concepts or methods are transformed to fit a different context from the original. The study of business history in Latin America, for instance, is conducted through the adaptation of the concepts of business history in high-income countries to low and middle-income countries. Later in its development that adaptation resulted in a differentiated discipline called Latin American business history.
- d) **Knowledge diffusion:** Knowledge diffusion occurs, for instance, when a concept that is not novel in mainstream JIS-indexed journals is introduced into a region and shared within the regional community. This can incentivise research on that subject in the regions, as was the case of the introduction of 16S ribosomal RNA sequencing to Colombia.
- e) **Teaching:** it is mainly the use of research in alternative JIS for teaching or learning-related activities, as in the introduction of PhD students to academic publications. For instance, papers in alternative JIS-indexed journals help to support and expand the content of their lectures. Lecturers use their own research in articles published in alternative JIS-indexed journals to teach students and use it in their modules.
- f) **Business model conversion:** this happens when a researcher publishes open access papers that incorporate bibliographic references from paid journals. For instance, in certain documents such as review papers researchers synthesise literature in mainstream JIS-indexed journals and make it available for readers that cannot afford access to mainstream JIS-indexed paid journals.
- g) **Multilingual referencing:** this is when researchers publish in non-English languages and incorporate references from journals in English and other languages. By referencing these multilingual sources researchers build on knowledge that can pose linguistic barriers to readers in their language. This is concretely seen in the incorporation of English-language references into research published in Spanish or Portuguese available through RedALyC and Scielo.

In certain occasions, a conjunction of the mechanisms above can stimulate new areas of study. For instance, the bulk of the production on Latin American business history is covered by alternative JIS, as was indicated by an interviewee and further confirmed through database queries. This sub-discipline emerged from personal interactions with American and British researchers on business history, and currently has grown into into a new area of study. It is mainly published in Spanish and the majority of papers circulate in alternative JIS-indexed journals. From this perspective, alternative JIS serve as a bridge to bring closer knowledge produced by perceived distant communities, with the potential to start novel avenues of research.

Knowledge-gap filling

Knowledge-gap filling is the coverage of knowledge that is neglected or not found in WoS. Examples from business and management, agricultural sciences, and chemistry showed that alternative JIS provide a space for the publication of distinctive original research.

- h) **Allow the publication of research that is not well covered or not found in WoS-indexed journals.** Examples include: research that is context-dependent such as Latin American business history or the conceptualisation and application of alternative indicators to understand innovation in countries with low patenting and R&D activity; distinctive subjects such as the production of passion fruit, and research on diseases affecting the production of oil palm; and certain disciplinary areas that have been displaced by others, such as the case of botany that has become less popular than pharmacognosy in high impact factor WoS-indexed journals.

The knowledge gaps that alternative JIS are fulfilling appear to be particularly important in all subjects in which local knowledge (“local” at various scales: from very localised to national to regional) is relevant for policy, management or industrial applications. For example, Arbeláez-Cortés (2013) documented that publications in alternative JIS play a major role in mapping Colombia’s biodiversity – an important topic given the country’s ecological wealth. In summary, alternative JIS offer a place for the publication of scientific knowledge beyond the boundaries of WoS and Scopus-indexed journals.

DISCUSSION AND CONCLUSIONS

In this article, we have examined the reasons why Colombian researchers publish in journals not indexed by mainstream JIS, i.e. WoS and Scopus. We have found that “lack of scientific quality” of their manuscripts is insufficient explanation to publication patterns. Instead, we have found that knowledge-gap filling, knowledge-bridging, and training towards mainstream JIS-indexed journals were the drivers for publishing in journals in alternative JIS. We believe that the reasons reported in Colombia are likely to apply to other countries in Ibero-America, as well as other developing countries. The extent to which they may also be relevant in other regions for certain topics and disciplines would need to be ascertained.

It follows from these findings that scientometric indicators based on mainstream JIS underrepresent some types of contributions of research from developing countries – as they do not value some training, knowledge-bridging and gap-filling functions. Therefore, research evaluations in Ibero-America should also consider publications in alternative JIS if they wish to value these other types of contributions, which may be particularly relevant in developing countries such as Colombia or other ‘peripheral’ contexts –i.e. in non-English contexts, for knowledge exchange with non-academic experts or for unconventional topics (Vessuri et al., 2014).

Besides the policy implications, the findings also relate to the theoretical discussion on the universalistic versus particularistic conceptualisations of science. Improvement of the scientific quality partly explains the *training* function of alternative JIS. Since lack of scientific quality is the perception of insufficient research competence as judged by global peers, this is a property that belongs to a universalistic conceptualisation of science. Hence, this publication behaviour can be partly explained by a Mertonian, universalistic conceptualisation of science as an institution.

However, researchers also publish in journals in alternative JIS in order to fulfil *knowledge-bridging* and *gap-filling* functions. Such publishing behaviour does not respond to a universalistic model of science, but to the recognition that mainstream JIS are particularistic institutions, with a lower coverage of journals from developing countries or non-English languages. Our findings thus support the view that scientific institutions such as bibliometric databases are located in specific contexts thus produce a representation of science from a specific, i.e. particularistic perspective. The value of using alternative JIS is to provide different particularistic perspectives of the scientific production, which may be valuable when the evaluation emphasis lies on situated and societal contributions of science (e.g. *gap-filling* and *knowledge-bridging*).

REFERENCES

- Arbeláez-Cortés, E. (2013). Knowledge of Colombian biodiversity: published and indexed. *Biodiversity and Conservation*, 22(12), 2875–2906. <http://doi.org/10.1007/s10531-013-0560-y>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <http://doi.org/10.1191/1478088706qp063oa>
- Garfield, E., 1996. Significant scientific literature appears in a small core of journals. *The Scientist*, 10(17), 13. Retrieved from <http://www.the-scientist.com/?articles.view/articleNo/18038/title/The-Significant-Scientific-Literature-Appears-In-A-Small-Core-Of-Journals/>
- Lemarchand, G. A. (2012). The long-term dynamics of co-authorship scientific networks: Iberoamerican countries (1973–2010). *Research Policy*, 41(2), 291–305. <http://doi.org/10.1016/j.respol.2011.10.009>
- Ordóñez-Matamoros, H. G., Cozzens, S. E., & Garcia, M. (2010). International Co-Authorship and Research Team Performance in Colombia. *Review of Policy Research*, 27(4), 415–431.
- Vessuri, H., Guédon J.-C. and Cetto, A.M. (2014) Excellence or quality? Impact of the current competition regime on science and scientific publishing in Latin America and its implications for development. *Current Sociology* 62(5), 647-665. <http://doi.org/10.1177/0011392113512839>



Identifying Sources of Scientific Knowledge: classifying non-source items in the WoS

Calero-Medina, C.M.

clara@cwts.leidenuniv.nl

Centre for Science and Technology Studies (CWTS), Leiden University, PO Box 905, Leiden, 2300AK (The Netherlands)

INTRODUCTION

The sources of scientific knowledge can be tracked using the references in scientific publications. For instance, the publications from the scientific journals covered by the Web of Science database (WoS) contain references to publications for which an indexed source record exist in the WoS (source items) or to references for which an indexed source record does not exist in the WoS (non-source items). The classification of the non-source items is the main objective of the work in progress presented here.

Some other scholars have classified and identified non-source items with different purposes (e.g. Butler & Visser (2006); Liseé, Larivière & Archambault (2008); Nerderhof, van Leeuwen & van Raan (2010); Hicks & Wang (2013); Boyack & Klavans (2014)). But these studies are focused in specific source types, fields or set of papers. The work presented here is much broader in terms of the number of publications, source types and fields.

DATA COLLECTION AND METHOD

The first step was to identify the non-source items collected by the WoS¹. In order to do so we just identified all the non-source items that appear on the references made by the articles and review articles published between 1980 and 2014 on the WoS. The set contains 297,904,154 distinct rows (the unique code number of the citing publication is included)

The information that appears per item in each paper may contains (it is not always the case) information at the level of *Author*, *Volume*, *Issue*, *Page Number*, and a string with *Other information* that may be filled with the title and/or the source. Table 1 shows some examples of non-source items in the WoS.

¹ The Web of Science (WoS) versions of the Science Citation Index and associated citation indices: the Science Citation Index (SCI), the Social Science Citation Index (SSCI), and the Arts & Humanities Citation Index (A&HCI); here the CWTS database containing these records as well as enhanced citation data is briefly indicated as CI. It is important to indicate that the conference proceedings database within the WoS database is not included in this study.

Table 1. Examples of Non-source Items in the WoS.

Author	Year	Volume	Page Number	Issue	Other Information
Fitzgerald, FS	1925				Great Gatsby
Descartes, R					Correspondance
Papavero, N	1978				Catalogue Diptera Am
*Austr Bur Stat	2001				Cens Pop Hous Cdata
Sorenson, DS	1995		1024		P 10 Ieee Puls Pow C
Ducange					Glossarium Mediae In
	1970	1	132		Encyclopedia Polymer
Caves, RE	1989	2	1225		Hdb Ind Org
Gui, MC					In Press J Therm Spr
*Off Nat Stat	2010				Ann Surv Hours Earn
Finney, GH	1975				Thesis Queens U King
Goodwillie, TG		3			Unpub Calculus
Perlez, J	1991		49		NY Times 0922
	1887		2		Figaro 0203
Puccini, G					Madama Butterfly

The non-source item has been identified and classified depending on the source type. As other studies have argued (Nerderhof, van Leeuwen & van Raan (2010) and Boyak & Klavans (2014)) the information at the level of Year, Volume, Issue and Page Number combined with the other fields can be used to help to estimate the type of source. For instance, the cases where all these four fields are filled may be considered Non-Source Journal or Non-Source Paper. The source type Non-Source Journal/Non-Source Paper has been the first in being identify, since it constitute the largest amount in the dataset. After that searching for keywords and key terms (in different languages) in *Other Information* has helped to identify Conference papers, Handbooks/Manuals, Thesis, Encyclopedia, Survey, In press, Preprint... In the case of Newspapers a more specific strategy has been followed looking for the main newspapers in different countries and then include the names (and possible variables) as search strategies. In the case of Reports (governmental and non-governmental reports) the information at the author level has been helpful since the non-governmental and governmental organizations are identified with an ‘*’.

In the case of the Books a semi-automatic process has been followed. We have combined the information at the level of *Other Information* and *Author* and search in fields where books are one of the main scientific output and therefore one of the main sources of knowledge as previous studies have shown (e.g. (Hicks (2009), Nerderhof, van Leeuwen & van Raan (2010)). A drawback of this approach is that Scientific and non-Scientific books have not been differentiated.

Table 2 shows how the Non-source Items from Table 1 have been classified following the method briefly explained above.

Table 1. Examples of Non-source Items in the WoS classified by Source Type

Author	Year	Volume	Page number	Iss	Other Information	Source Type
Fitzgerald, FS	1925				Great Gatsby	Book
Descartes, R					Correspondance	Book
Papavero, N	1978				Catalogue Diptera Am	Catalogue
*Austr Bur Stat	2001				Cens Pop Hous Cdata	Census
Sorenson, DS	1995		1024		P 10 Ieee Puls Pow C	Conference Paper
Ducange					Glossarium Mediae In	Glossary
	1970	1	132		Encyclopedia Polymer	Encyclopedia
Caves, RE	1989	2	1225		Hdb Ind Org	Handbook
Gui, MC					In Press J Therm Spr	IN PRESS
*Off Nat Stat	2010				Ann Surv Hours Earn	Survey
Finney, GH	1975				Thesis Queens U King	Thesis
Goodwillie, TG		3			Unpub Calculus	Unpublished
Perlez, J	1991		49		Ny Times 0922	Newspaper
	1887		2		Figaro 0203	Newspaper
Puccini, G					Madama Butterfly	Music

PRELIMINARY RESULTS AND FOLLOW UP RESEARCH

Around two thirds of the initial data non-source dataset have been classified. Overall we have identified 44 source types (see Table 3 below). The frequency of appearances varies greatly but having a refine source type will be of great help for future analysis. The most frequent ones are Non-Source Journals and Non-Source Papers. Under this category are many papers published in journals that are actually covered by the WoS but they are from volumes previous to 1980² and papers published after 1980 in Journals only partially covered by the WoS. Conference Papers and Books are also very frequent. Newspapers and Magazines constitute also quite frequent types.

² The study is based on the WoS database with publications from 1980 onwards.

Table 3. Source Types Identified

• Archive	• Magazine
• Blog	• Meeting
• Book	• Monograph
• Case Report	• Music Piece
• Catalogue	• Newsletter
• Cd Rom	• Newspaper
• Census	• Non-Source Journal/Non-Source Paper
• Cited indirectly	• Performance
• Communication(*)	• Picture/Illustration/Painting
• Conference Proceedings	• Preprints
• Dictionary/Vocabulary/Thesaurus/Lexicon/Glossary	• Report
• Documentary	• Seminar/Workshop
• Ejournal	• Statistics
• Encyclopedia	• Survey
• Festival Related	• Thesis
• Film/Movie	• Twitter
• Gazette	• Unpublished
• Handbook/Manual	• Video
• In Press	• Website
• Int Tables Cryst	• Wikipedia
• Interview	• Working Paper
• Journal_Periodical	• Youtube

The two main next steps previous to create the final version of the Non-Source Database are:

- For each of the Source Types, select a sample to check the validity of the Source type assignments in order to provide reliable estimates on the validity of our assessments. This is especially important for the Conference Papers, Books and Reports.
- Select a sample from the data that could not be classified yet to learn if some of them could be classified.

Additionally, there will be an attempt to reclassify books in scientific and non-scientific using some mapping and clustering techniques with the help of the VOSviewer software (van Eck & Waltman, 2010).

REFERENCES

- Butler, L. & Visser, M.S. (2006). Extending citation analysis to non-source items. *Scientometrics*, 66, 327-343.
- Hicks, D. (1999). The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. *Scientometrics*, 44(2), 193-215.
- Hicks, D. & Wang, J. (2013). The *New York Times* as a Resource for Mode 2. *Science, Technology & Human Values*, 38(6), 851-877.
- Lisée, C., Larivière, V. & Archambault, E. (2008). Conference Proceeding as a Source of Scientific Information: A Bibliometric Analysis. *Journal of the American Society for Information Science and Technology*, 59(11), 1776-1784.
- Nederhof, A.J., van Leeuwen, T.N. & van Raan, A.F.J. (2010). Highly cited non-journal publications in political science, economics, and psychology: A first exploration. *Scientometrics*, 83, 363-374.
- Van Eck, N.J. & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523-538.



Data quality and consistency in Scopus and Web of Science in their indexing of Czech Journals

Pavel Mika,* Jakub Szarzec** and Gunnar Sivertsen***

* mika@knav.cz

Library of the Czech Academy of Sciences,
Národní 1009/3, 115 22 Praha 1, Czech Republic

** jakub.szarzec@techlib.cz

National Library of Technology,
Technická 6/2710, 160 80 Praha 6 – Dejvice, Czech Republic

*** gunnar.sivertsen@nifu.no

Nordic Institute for Studies in Innovation, Research and Education (NIFU)
P.O. Box 2815 Tøyen, N-0608 Oslo, Norway

ABSTRACT

This study addresses the discussion of “quality versus coverage” that often arises if a choice is needed between Scopus and Web of Science (WoS). We present a new methodology to detect problems in the quality of indexing procedures. Our preliminary findings indicate the same degree and types of errors in Scopus and WoS. The more serious errors seem to occur in the indexing of cited references, not in the recording of traditional metadata.

Keywords

Scopus; Web of Science; data quality; journal coverage; citation indexing; references; Levenshtein distance.

Submission type: Research in progress paper.

Relevant track: Data infrastructure for research metrics.

INTRODUCTION

This study addresses the discussion of “quality versus coverage” that often arises if a choice is needed between Scopus and Web of Science (WoS). With regard to coverage of source documents and citations, there are large differences in favour of Scopus, although there is not full overlap with WoS content (Gavel & Iselid 2008). The consequences of different coverage depend on the purpose of a particular usage. The two data sources need to supplement each other from an information retrieval perspective (Bar-Ilan 2010). They can, however, replace each other as the basis for indicators of scientific production and citations at the country level (Archambault et al. 2009), but less so at the level of institutions (Vieira & Gomes 2009) or in fields of research that tend to be marginally covered in both sources (Bartol et al. 2014; Haddow & Genoni 2010; Sivertsen 2014).

The quality and consistency of citation indexing procedures are important for all purposes, however. Franceschini et al. (2015) recently published indications of serious types of errors in Scopus that WoS is not free from either. Our study aims at resolving the same question of

data quality. We present a new methodology to reach this aim. Our preliminary findings indicate fewer errors and less difference in this respect between Scopus and WoS than we expected from the earlier study. More serious errors seem to occur in the indexing of cited references, not in the recording of traditional metadata. Our further research – also to be presented at the conference – will clarify the extent of this problem.

METHODS

We chose to study journals published by organizations or publishers in the Czech Republic. The reason for this choice is that we wanted to compare Scopus and WoS mainly where they differ: in coverage of the “periphery” of the international core journals. We chose the Czech Republic because the printed versions of the indexed journals are easily available to us. There are 49 Czech journals in the 2014 edition of the Journal Citation Report (WoS) and 159 Czech journals in the 2014 Scopus Journal Title List. Among these, 46 journals are indexed in both databases. They cover Agriculture, Chemistry, Business Economics, Engineering, Plant Sciences, Food Science Technology, Veterinary Sciences, Entomology, Physiology and Microbiology. Most of them (84 per cent) are published in the English language; some are bilingual; the remaining few publish in the Czech language only.

We downloaded the data manually in early December 2015 using the web interface of each database. The queries were limited by ISSN for five years, 2010-2014. We retrieved 13,281 records from Scopus and 13,947 records from WoS in the same 46 journals. The completeness of both downloads was checked against the online versions of the databases after download.

Matching supposedly identical records was crucial in the preparation of data for further analysis. We used an iterative process in several phases where we combined manual and automatic methods based on the Levenshtein distance metric. We were able to match a total of 12,494 records. The matched records thereby constituted 94 percent of the records retrieved from Scopus and 90 percent of the records retrieved from WoS.

The quality and consistency of the data in the two databases was studied by making two types of systematic comparisons. First, the matched records were compared to each other to study possible differences in indexing between the two databases. Second, all records, including those that could not be matched, were compared to the electronic archives of the indexed journals. In addition, two of the journals were analysed using their printed versions. In both types of comparisons, the official indexing policies of the two databases (Scopus Elsevier 2016; Thomson Reuters 2016), which are not identical, provided important guidelines with regard to expected outcomes.

RESULTS

The results of the comparison of the 12,494 matched records are shown in Table 1.

Table 1. Comparison of selected fields

WoS abreviation/name of field	Scopus name of field	Number of identical (provided) values	Base for percentage	% of identical (provided)	Comparison method
AU Authors	Authors	12,405	12,494	99.3	Number
TI Document Title	Title	8,394	12,383	67.8	Levensthein
DT Document Type	Document Type	11,424	12,318	92.7	Identical YES/NO
TC Times Cited	Cited by	3,713	12,494	29.7	Number
PY Year Published	Year	12,452	12,494	99.7	Identical YES/NO
VL Volume	Volume	12,325	12,494	98.4	Identical YES/NO
IS Issue	Issue	11,766	12,494	94.1	Identical YES/NO
BP Beginning Page	Page start	12,302	12,494	98.4	Identical YES/NO
EP Ending Page	Page end	11,944	12,494	95.6	Identical YES/NO
DI Digital Object Identifier (DOI)	DOI	2,235	2,296	97.3	Identical YES/NO
LA Language	Language of Original Doc.	11,186	12,494	89.5	Identical YES/NO
DE Author Keywords	Author Keywords	12,015	12,494	96.1	Number
AB Abstract	Abstract	11,901	12,494	95.3	Provided YES/NO
NR Cited Reference Count	Reference count	3,376	4,445	76.0	Number

Generally, we find a high degree of consistency in indexing between the two databases, measured as the percentage identical data in each field, with one important exception, the number of references. All smaller or larger differences between the two databases can be technically explained without altering the general impression that the metadata are of relatively high quality in both databases. Here are several explanations before we turn a discussion of the exception:

- A higher rate of identical titles (68%) could not be expected, because 20 percent of the Scopus titles are multilingual. Other differences were caused mainly by the transcription of technical terms using the Greek alphabet into Latin, for Scopus titles.
- The number of times cited is expected to be different because the two databases cover different numbers of source journals.
- The differences in document type classification are mainly explained because the two resources use different classification schemes. The differences are small. The most common differences are shown in Table 2.

Table 2. Document type differences

WoS doc. type	Scopus doc. type	Number of docs.	% of explored dataset
Article	Review	208	1.7
Review	Article	205	1.7
Proceedings Paper	Article	171	1.4
Editorial Material	Article	139	1.1
Article	Proceedings Paper	107	0.9
Editorial Material	Review	17	0.1

Document type information is important in bibliometric analysis in order to normalize citation indicators. Our results indicate that this type of information is relatively reliable. However, even more important is the indexing of the reference lists in each document. An exception to the finding that metadata are of high quality is the indication we get as we see that 24 percent of the matched records have different reference counts in Scopus and WoS. This is a clear indication that the reference lists in the source documents are not appropriately or fully indexed.

We found 222 WoS records with more references than in Scopus and 847 Scopus records with more references than in WoS. The number of missing references for each comparison is shown in Table 3. The most common difference (12%) was caused by one missing reference in WoS records.

Table 3. Differences in number of references

Reference difference	Number of records	% of records	Number of missing references
WoS>SC	222	5	-1,913
SC>WoS	847	19	2,005
SUM	1,069	24	92

This observation of differences was the starting point for further research when we tried to compare all references from observed records. Unfortunately we still weren't able to match all the references to find out any pattern in missing (or excess) references.

In the second part of the study, we compared matched as well as unmatched records (Scopus versus WoS) to the electronic archives of the 46 indexed journals. A total of 17,759 records could be used for the study of how and to what extent the journals are indexed. A quantitative overview is given for each of the journals in Table 4 (Appendix). Here, we compare the

number of records in the original source journal to the number of records indexed in WoS and Scopus and the number of records that could be matched between them. No numbers are the same for any of the journals and there are wide differences for some journals. The right column in Table 4 (A-C) refers to the following explanations for the differences:

- A. There are only small differences for nine journals. The differences can mainly be explained because of differently defined document types used for indexing hybrid journals with a large array of document types.
- B. There are larger differences between Scopus and WoS for 25 journals; however, the number in one of the databases resembles the number of records in the original source. The differences between the two indexing databases can be explained by differing indexing policies, with the exceptions below.
- C. There are large differences between the original sources and the two indexing databases for nine journals. In these cases, we found that the electronic archive of the journal does not cover the journal completely or the archive includes supplemental items not published in the regular journal.

An example of C is *Chemické listy* (0009-2788), where the archive includes supplementary material such as conference abstracts of plenary lectures, oral sessions and posters.

Differences of type B were examined by inspecting the printed versions of two journals. In *Folia Biologica* (ISSN 0015-5500), we discovered that the larger number of records in Scopus was caused by an error in which 71 records from a Polish journal with the same name but different ISSN (0015-5497) were included. We also found two instances of duplicate records in Scopus. All in all, we found 14 cases of the duplicate Scopus records in the whole dataset, which is less than expected from earlier studies of the same error (Valderrama-Zurián et al. 2015).

Inspecting *Československá psychologie* (0009-062X) in the same way, we found that neither Scopus nor WoS covered this journal completely. In spite of the indexing policy, 12 items were not indexed in WoS – mostly news, errata, and discussions. Of 214 items not indexed by Scopus, 51 were classified as research articles in WoS. If this classification is correct, they should have been indexed in Scopus according to its policy. The other missing items in Scopus can be explained by the policy of not indexing such items.

DISCUSSION AND FOCUS FOR FURTHER RESEARCH

We have established a methodology for two types of comparisons that aim to test the quality and consistency of the data and indexing in Scopus and WoS, by:

- a. Matching and measuring the degree of similarity in supposedly identical records in both databases.
- b. Comparing data from both databases to the sources that were indexed.

With both methods, most of the differences we observed could be explained according to differing methods and policies for indexing in Scopus and WoS or the specific publishing policies of journals.

There are two major exceptions, however, that will be the focus of our further studies:

- a. Differences in the number of cited references in a record may be an indication that reference lists in the source documents are not appropriately or fully indexed.
- b. Differences between the number of records in the archive of the source journal and the databases can be an indication that the contents are not appropriately or fully indexed.

REFERENCES

Archambault, É, Campbell, D., Gingras, Y., & Larivière, V. (2009). Comparing bibliometric statistics obtained from the Web of Science and Scopus. *Journal of the American Society for Information Science and Technology*, 60(7), 1320-1326.

Bar-Ilan, J. (2010). Citations to the “Introduction to informetrics” indexed by WOS, Scopus and Google Scholar. *Scientometrics*, 82(3), 495-506.

Bartol, T., Budimir, G., Dekleva-Smrekar, D., Pusnik, M., & Juznic, P. (2013). Assessment of research fields in Scopus and Web of Science in the view of national research evaluation in Slovenia. *Scientometrics*, 98(2), 1491-1504.

Franceschini, F., Maisano, D., & Mastrogiacomo, L. (2016). The museum of errors/horrors in Scopus. *Journal of Informetrics*, 10(1), 174-182.

Gavel, Y., & Iselid, L. (2008). Web of Science and Scopus: A journal title overlap study. *Online Information Review*, 32(1), 8-21.

Haddow, G., & Genoni, P. (2010). Citation analysis and peer ranking of Australian social science journals. *Scientometrics*, 85(2), 471-487.

Scopus Elsevier. (2016). Content Policy and Selection. (<https://www.elsevier.com/solutions/scopus/content/content-policy-and-selection>) [21 March 2016].

Sivertsen, G. (2014). Scholarly publication patterns in the social sciences and humanities and their coverage in Scopus and Web of Science. In E. Noyons (Ed.), *Proceedings of the science and technology indicators conference 2014 Leiden* (pp. 598-604). Leiden: CWTS.

Thomson Reuters. (2016). The Thomson Reuters Journal Selection Process. (<http://wokinfo.com/essays/journal-selection-process/>) [21 March 2016].

Valderrama-Zurián, J., Aguilar-Moya, R., Melero-Fuentes, D., & Aleixandre-Benavent, R. (2015). A systematic analysis of duplicate records in Scopus. *Journal of Informetrics*, 9(3), 570-576.

Vieira, E. S., & Gomes, J. A. (2009). A comparison of Scopus and Web of Science for a typical university. *Scientometrics*, 81(2), 587-600.

Table 4. Number of records in the original source journal compared to the number of records indexed in WoS and Scopus, and the number of records that could be matched between them.

The right column (A-D) refers to explanations for the differences given in in the text.

Journal title abbrev.	Source N	WoS N	SC N	Matched N	Differences
Acta Ent Mus Nat Pra	293	291	241	239	B
Acta Geodyn Geomater	216	218	217	217	A
Acta Chir Orthop Tr	393	346	360	340	C
Acta Vet Brno	388	398	388	388	B
Agr Econ-Czech	304	302	292	292	B
Appl Math-Czech	176	175	169	169	B
Biomed Pap	728	346	344	317	C
B Geosci	234	233	245	220	B
Cent Eur J Publ Heal	233	224	286	219	B
Ceram-Silikaty	278	280	278	278	A
Cesk Slov Neurol N	668	594	574	536	C
Cesk Psychol	223	378	176	171	B
Czech J Anim Sci	331	320	317	317	C
Czech J Food Sci	408	406	386	386	B
Czech J Genet Plant	208	207	192	191	B
E M Ekon Manag	237	286	234	232	B
Epidemiol Mikrobi Im	193	171	165	132	C
Eur J Entomol	445	409	403	398	C
Financ Uver	136	137	136	130	A
Folia Biol-Prague	199	198	272	198	B
Folia Geobot	144	136	137	135	C
Folia Microbiol	439	439	457	437	B
Folia Parasit	237	232	211	211	B
Folia Zool	198	193	194	191	A
Fottea	106	108	103	103	A
Hortic Sci	128	128	126	126	A
Chem Listy	4,160	1,290	1,254	1,033	C
J Appl Biomed	133	129	106	103	B
J Geosci-Czech	127	126	111	110	B
Kybernetika	344	343	337	335	B
Listy Cukrov Repar	603	416	456	390	C
Morav Geogr Rep	109	86	108	84	B
Neural Netw World	218	212	205	198	A
Photosynthetica	379	373	401	364	B
Physiol Res	631	634	619	610	B
Plant Protect Sci	134	97	132	88	B
Plant Soil Environ	445	437	437	437	A
Polit Ekon	266	279	218	214	B
Prague Econ Pap	136	138	128	127	B
Preslia	127	127	127	120	A
Radioengineering	728	735	725	724	A
Slovo Slovesnost	177	160	73	66	B
Sociol Cas	760	471	204	197	C
Soil Water Res	111	110	111	110	A
Stud Geophys Geod	224	224	232	223	B
Vet Med-Czech	404	405	394	387	B



Missing citations due to exact reference matching: Analysis of a random sample from WoS. Are publications from peripheral countries disadvantaged?¹

Paul Donner

donner@dzhw.eu

Deutsches Zentrum für Hochschul- und Wissenschaftsforschung, Berlin
Schützenstraße 6a, 10117 Berlin (Germany)

INTRODUCTION

Citation counts of scientific research contributions are one fundamental data in scientometrics. Accuracy and completeness of citation links are therefore crucial data quality issues (Moed, 2005, Ch. 13). However, despite the known flaws of reference matching algorithms, usually no attempts are made to incorporate uncertainty about citation counts into indicators. This study is a step towards that goal. Particular attention is paid to the question whether publications from countries not using basic Latin script are differently affected by missed citations.

The proprietary reference matching procedure of Web of Science (WoS) is based on (near) exact agreement of cited reference data (normalized during processing) to the target papers bibliographical data. Consequently, the procedure has near-optimal precision but incomplete recall - it is known to miss some slightly inaccurate reference links (Olensky, 2015). However, there has been no attempt so far to estimate the rate of missed citations by a principled method for a random sample. For this study a simple random sample of WoS source papers was drawn and it was attempted to find all reference strings of WoS indexed documents that refer to them, in particular inexact matches. The objective is to give a statistical estimate of the proportion of missed citations and to describe the relationship of the number of found citations to the number of missed citations, i.e. the conditional error distribution. The empirical error distribution is statistically analyzed and modelled.

DATA AND METHODS

The analyzed data originate from licensed raw data in tagged format of the WoS journal and proceedings citation indexes. A simple random sample of target items was drawn from all journal articles, letters and reviews, as defined in the data. The WoS citation links were obtained, as given by the WoS matchkey, the T9/R9 fields. All citations until 2015 are counted.

Reference data of all publications from 1980 to spring 2015 are indexed for search, the sampled target source items are from the same period. The reference strings, consisting of author name field, split into last name and initials at the comma, the source title, publication year and first page fields were indexed with Oracle Text. The volume field was not considered because for the target journal items volume and publication year are nearly redundant information but publication year is more accurate and more complete in WoS

¹ This work was supported by BMBF project 01PQ13001. The author would like to thank Anastasiia Tcypina for her help in collecting the data.

reference data than volume (Moed, 2005, table 13.1). A procedure was programmed to search the index for references likely referring to the sampled target items. Because the search field data of the target items has to correspond to the way the reference data is prepared by WoS, the target author name and source title were pre-processed accordingly. This entailed reduction of first name to initials, the removal of non-letter characters in the name, and, for the source title, using the WoS abbreviation. Where more than one abbreviated title version for the same journal existed in the data and when an additional group first author was available, all possible combinations of those fields' values were used as search input. The procedure performs a fuzzy search on the index and returns a list of unique candidate reference strings that are sufficiently similar to the target input. The search is deliberately lenient so that all possible matches are returned in order to prevent false negatives as much as possible, which is a requirement for this study.

The candidates were reviewed clerically on whether they constitute a match to the target or not by a student assistant. Ambiguous candidates were afterwards assessed by the author. Care was taken to avoid false positive matches by querying the database for any exact matches of the candidate reference strings other than the target item. The found positive matches are used as additional citation links and the derived extended citation count for each target item is calculated by retrieving all references using those candidate strings.

PRELIMINARY RESULTS

A total of 372 cases were assessed. The distribution of missed citations per item is presented in table 1.

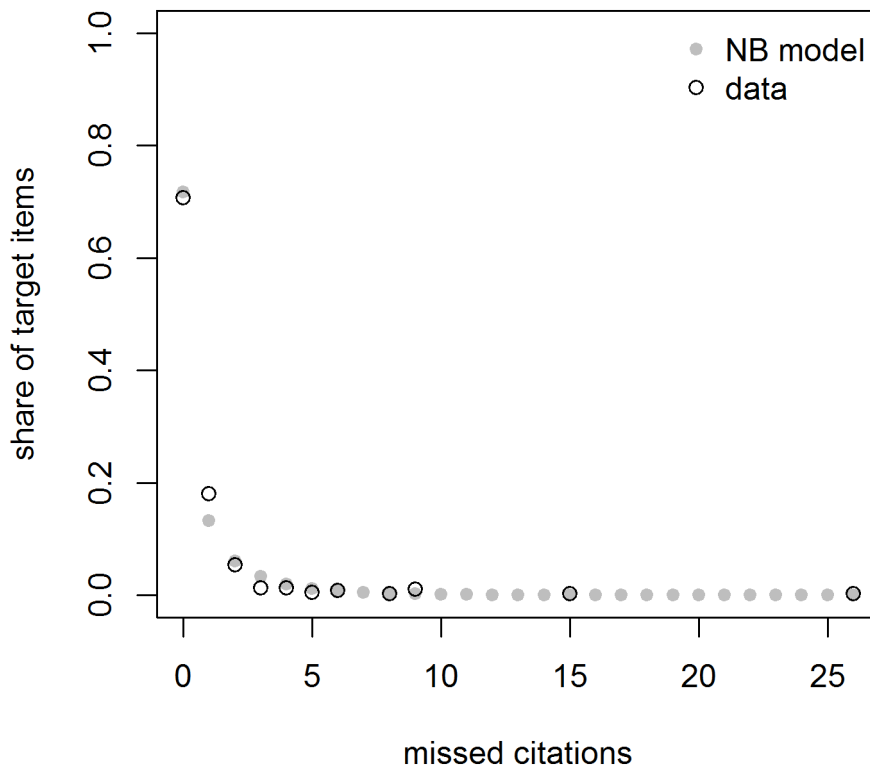
Table 1: Distribution of number of missed citations

missed citations	0	1	2	3	4	5	6	8	9	15	26
occurrence	263	67	20	5	5	2	3	1	4	1	1

29.4 % of the target items have one or more detected missing citation. The average of citations per paper (CPP) according to WoS is 16.4; the average of missing CPP is 0.7. In total, according to WoS' exact citation links, the papers were cited 6120 times. 255 additional citations were found. Thus, 4 % of citations were missed. An association between the apparent citation count and the citation count error can be observed, as the number of WoS citations and the number of missed citation per item are correlated with $r = 0.31$.

Citation distributions can be approximately modelled by negative binomial (NB) distributions with reasonable accuracy (Schubert and Glänzel, 1983; Ajiferuke and Famoye, 2015). As the error distribution is also discrete, non-negative and heavily skewed, it was attempted to model it with the NB distribution. Numerical estimation using the R package *MASS* gave the following parameter estimates: $\theta = 0.254$ (SE = 0.039) and $\mu = 0.685$ (SE = 0.083) and the model fit is depicted in Figure 1.

Figure 1: Empirical probability distribution of the data and first NB model



The goal is not just to describe the distribution in general by a model, but to predict the error, that is, the number of missing citations, for a given number of WoS citations. The above model does not take the association between WoS citation count and missed citations into account, which was shown by the correlation. The model is next extended by regressing the parameter μ on the WoS recorded citation counts w , while holding the scale parameter θ of the binomial regression fixed to the previously estimated value.

The estimated equation is: $\mu = 0.5 + 1.012 \times w$.²

Having estimated the parameters, we can now simulate random deviates from this distribution or obtain values of the density or cumulative probability at any desired point. For example, according to the model, the probability of having zero missing citations for a publication with WoS citation count of 0 is 75.9 %, for one missing citation is 12.8 %. For 100 WoS citations, the probability of having 0 missing citations is 21.8 %; the probability for one missing citation is 5.5 %.

By Monte Carlo simulation from the model one may obtain a predicted *distribution* of the sum of missed citations, in this case for 372 publications. In a Bayesian statistical framework this is the posterior probability distribution of the parameter of interest. To do this, we make 372 draws from the model, that is, a random NB variable with the estimated model

² SE of the intercept: 0.182. SE of the coefficient of w : 0.004; both significant at the 0.01 level.

parameters. To take model uncertainty into account, inputs (the θ and μ parameters) are not fixed but are also randomly drawn from truncated normal distributions (as only values > 0 are possible) with the mean being the point estimate and the standard deviation being the estimated standard error. This means that parameter estimates are replaced by prior distributions with hyperparameters from the preceding estimations.

The sum of the values of an iteration is the estimate of the sum of missed citations. The procedure was replicated 10.000 times to get an approximation of the probability distribution of the number of missed citations, given the model and estimates. The distribution obtained is characterized in table 2:

Table 2: Summary of the simulated distribution of the sum of errors (10.000 replications)

min.	median	mean	max.	credible 95 % interval
131	257	260	492	(195, 341)

Recall that the observed total of missing citations is 255. The model point estimate of missed citations is 260 with a Bayesian credible interval derived from the quantiles of the posterior distribution of [195, 341].

PERIPHERAL EFFECTS

Eastern European and East Asian researchers often encounter ambiguities when they have to transliterate their names in to basic Latin script or when their names are simplified to basic Latin script for database indexing. Because reference matching relies on author names, it is hypothesized that publications from those peripheral regions are subjected to comparatively higher risks of missed citations. To test the hypothesis, all first author country information were coded into three mutually exclusive regions and one category for unknown country. This nominal variable with three levels was added to the second model. Three papers' first authors had both one address of a peripheral region and one of a non-peripheral one. In these cases, they were coded as the peripheral region. The distribution of publications over regions is displayed in table 3. Furthermore, publication year was also added to the model as a predictor to see if any temporal change in reference accuracy can be detected.

Table 3: Regions defined as "peripheral" and countries

Region	Publications	Countries
East Asia	48	China, Japan, South Korea, Taiwan
Eastern Europe	23	Ukraine, Russia, USSR, Serbia, Bulgaria, Czech Republic, Czechoslovakia, Serbia and Montenegro, Hungary, Poland, Latvia
Others	301	all others

The regression equation is

$$\mu = \beta_1 + \beta_2 w + \beta_2 r + \beta_3 p;$$

with r being the variable region, p the publication year and the parameter θ held constant to 0.254 as before.

In the expanded regression, the coefficients for the regions were found to be not significant at the 0.05 level with 'Others' as the reference level. The effect of publication year was not

significant either³. The coefficient of WoS citation count (β_2) is slightly smaller but remains significant at the 0.01 level. Thus, the hypotheses that the country of the first author or the publication year affects the citation error rate are rejected for this sample. The clear limitation of this study is that the group sizes are so small that differences are difficult to detect, so an extended sample, possibly stratified by region, might reveal contradictory evidence.

References

- Ajiferuke, I., & Famoye, F. (2015). Modelling count response variables in informetric studies: Comparison among count, linear, and lognormal regression models. *Journal of Informetrics*, 9(3), 499-513.
- Moed, H. F. (2005). *Citation Analysis in Research Evaluation*. Springer, Dordrecht, The Netherlands.
- Olensky, M. (2015). *Data accuracy in bibliometric data sources and its impact on citation matching*. Dissertation, Humboldt-Universität zu Berlin.
- Schubert, A., & Glänzel, W. (1983). Statistical reliability of comparisons based on the citation impact of scientific publications. *Scientometrics*, 5(1), 59-73.

³ The publication year was significant at the 0.1 level.



Funding Acknowledgements in the Web of Science: inconsistencies in data collection and standardization of funding organizations¹

Jeroen van Honk^{*}, Clara Calero-Medina^{*} and Rodrigo Costas^{*}

[*j.s.van.honk@cwts.leidenuniv.nl](mailto:j.s.van.honk@cwts.leidenuniv.nl); clara@cwts.leidenuniv.nl; rcostas@cwts.leidenuniv.nl

Leiden University (Centre for Science and Technology Studies - CWTS), P.O. Box 905, Leiden, 2300 AX (The Netherlands)

ABSTRACT

Funding acknowledgements (FAs), as provided by the Web of Science, are a recent inclusion in the bibliometric toolset. They are starting to be used to study the presence, relationships and effects of funding and funders in the production of new scientific research. The incorporation of this new source of information comes with the need to understand how these data are collected and indexed in the database. This paper discusses important inconsistencies related to the method by which the data for FA and funders are selected, extracted and indexed by WoS, thereby highlighting the need to construct a thesaurus for the data. Problematic areas are found to be the quality of the input data and the conceptualization of what does and does not constitute a funding acknowledgement. Based on this critical analysis of the data and the identification of the main issues, we provide several recommendations for researchers, funders, WoS and other users of the data.

INTRODUCTION

Funding acknowledgements (FAs) have recently started to be included in the bibliometric toolset as a source of information to study the presence, relationships and effects of funding and funders in the production of new scientific research. This has been possible particularly since the Thomson Reuters Web of Science (WoS) database started to collect FA information from August 2008 onwards. The inclusion of this relevant piece of information has opened up new possibilities in the field of acknowledgements research (Costas & van Leeuwen, 2012; Desrochers et al., 2015; Díaz-faes & Bordons, 2014) and particularly in the area of FA studies (Sirtes, 2013). The availability of FAs in scientific publications allows the study of the presence of funding across disciplines, different funders or co-funding patterns in science (Wang & Shapira, 2011). It is important to highlight that in addition to the FA information, Thomson Reuters also collects the full funding text from scientific publications as well as the grant number, if provided in the publication, thus opening the possibility of more refined analysis of specific funding programs.

However, the incorporation of this new source of information also comes with the need to understand how these data are collected and indexed in the database. The importance of knowing the boundaries of the data collected is critical for the adequate use of this new source of information with analytical purposes (Paul-Hus et al, 2016). Previous studies have already pointed out some of the limitations of the FA data collected by WoS. For example, Rigby (2011) pointed out the presence of misspellings of funding bodies or errors in grant numbers, a problem that has also been addressed by Sirtes (2013), which sought to correct the severe

¹ This work has been partially supported by funding from the DST-NRF Centre of Excellence in Scientometrics and STI Policy (South Africa).

problems of data standardization in WoS funding information. Additionally, important limitations on the coverage of these data have recently been reported (Paul-Hus et al, 2016). In Paul-Hus et al (2016) an internal guideline policy document from Thomson Reuters was discussed (Thomson Reuters Bibliographic Policy Funding Acknowledgements, 2015). The existence of these guidelines supports the idea that the FA data collection has an important decision component, where indexers are expected to evaluate and select funding texts and funding bodies for their indexation (or not), thus opening also possibilities of introducing inconsistencies in the selection and indexation of the data.

OBJECTIVES

This paper highlights the challenges that arise in the construction of a thesaurus of funding organizations based on the funding acknowledgement data from the WoS. We will discuss important inconsistencies related to the method by which the data for FA and funders are selected, extracted and indexed by WoS. Thus, the main objective of this paper is to provide a critical analysis on central methodological aspects related with the collection and standardization of FA data in WoS.

METHODOLOGY AND DATA COLLECTION

The WoS source data are collected on three levels, namely (1) the full funding text, (2) the extracted funding body, and (3) the extracted grant number. The funding body and grant number are linked where they occur together.

A cursory look at these funding bodies indexed by WoS makes it abundantly clear that, without any standardization, the quality of data is highly variable. It can be argued that there are two main central issues related to FA data quality: 1) great variation in the acknowledging practices held up by researchers, thus creating a diversity of funding organization names², grant numbers, mentions of support, etc.; and 2) the relatively undisclosed and occasionally unclear means by which WoS identifies and collects the FA data. These two obstacles together make the data challenging and specific solutions (e.g. thesauri, contextual analysis of the funding texts) need to be considered in order to develop meaningful analyses from the FA data. The two main issues we encountered in the standardization of FA WoS data can be grouped under: 1) inconsistencies in the selection of FA data to be indexed; and 2) inconsistencies of the FA data indexed in WoS.

Inconsistencies in the selection of FA data to be indexed

It has been reported that WoS only collects FAs when the acknowledgement³ section of the publication contains funding-related information (cf. Paul-Hus et al, 2016). In our database, barring a negligible number of exceptions⁴, there are no funding texts which are not connected to either at least one funding body or grant number. However, the criteria regarding which types of funding information are selected and indexed for which acknowledgements remain unclear. An entity can be indexed as funding body after simply being thanked for

² These are caused, among many other things, by spelling mistakes, varying translations, and formatting variances.

³ Acknowledgement sections generally contain more than just funding acknowledgements, with acknowledging practices tending to extend to more generic expressions of support (Costas & van Leeuwen, 2012; Díaz-faes & Bordons, 2014).

⁴ A total of only 104 cases, most of them from 2009, when the data were just starting to be collected by WoS.

“support”, without further specification as to the form this support took.⁵ Whether such unspecified cases actually concern financial, technical, intellectual, or material support is not always evident, even when taking into account the rest of the funding text.

Those funding bodies that are extracted are not standardized. The method seems to be restricted to a simple tagging of phrases occurring in the text. While such a method seems intuitive enough, it proves to be problematic in its distinction between implicit and explicit funding acknowledgements. Consider the following pair of examples:

“This study was supported by internal funding from UNC Health Care”.

“This study was supported by internal funding sources”.

Of these two, the first text is indexed, with “UNC Health Care” listed as funding body. The second is not indexed. In order to be consistent, these cases ought to be treated equivalently. Another example can be the following:

“The authors wish to thank the National Institute of Malaria Research (Indian Council of Medical Research), Delhi, India for encouragement and moral support.”.

It is unclear why this acknowledgement is indexed (and the “National Institute of Malaria Research (Indian Council of Medical Research), Delhi, India” extracted as funder). Obviously this is not a FA and the support mentioned is “moral”, not financial.

It is important to remark that the study of this type of inconsistency is very complex as it requires an analysis of which acknowledgements have been selected for indexing and which have not. We are currently working on a more extensive analysis to explore further these limitations (van Honk, Calero-Medina, & Costas, 2016).

Inconsistencies in the indexation of funding bodies

Like the decision of what constitutes a funding text, the extraction of funding bodies from these texts is not always self-evident. One problem arising in the WoS FA data is when two funding entities have been incorrectly lumped together and presented as one (i.e. “NSF/DOE” and “National Science Foundation/Department of Energy”, 346 occurrences in the WoS data). The reverse also occurs: two funding agencies identified separately while they are actually parts of the same whole. This happens for instance with the “Program for Changjiang Scholars and Innovative Research Team in University (PCSIRT)” string, where the part before and after “and” are occasionally (yet not consistently) indexed separately. The lack of consistency is notable: in otherwise similar circumstances one acknowledgement is split while the other is kept whole.⁶ A similar inconsistency is found for the “U.S. EPA’s Science to Achieve Results” scholarship, which is sometimes split in “U.S. EPA’s Science” and “Achieve Results” (this happened in at least 85 cases). These examples suggest that a manual examination on a paper-by-paper basis is bound to introduce inconsistencies in the data

⁵ i.e. “BP would like to thank ICRA Net for support on this project.” ICRA Net is in fact an Italian research institute, and this unqualified “support” could have taken many forms, yet the institute is indexed as a funding body by WoS.

⁶ Within acknowledgement texts containing the phrase “Changjiang Scholars and Innovative Research Team in University”, there are many more indexations for the second part of the conjunction (474) than for the first (86).

extraction and parsing, especially when what constitutes a “funding acknowledgement” seems to have been loosely conceptualized by WoS.

The clue to another important inconsistency is in the occurrence as funding bodies of organizations that are not strictly funding organizations (e.g. universities acknowledged and indexed by WoS as funders). It turns out that when universities occur as FA they usually occur as the recipient of funding rather than as funder. Take for instance the following funding text:

“This material is based upon work supported by the National Science Foundation under Grants Nos. CHE-0721505 and CHE-0809053 at the University of Arizona (SGK) and under Grant No. CHE-719157 at the Massachusetts Institute of Technology. Cossairt and Cummins would also like to thank Thermphos International for support.”

From this text, WoS has extracted three funders: “National Science Foundation at the University of Arizona”, “Massachusetts Institute of Technology” and “Thermphos International”. Here, two universities are indexed as funding bodies while in reality only one funding body is mentioned here: the National Science Foundation.⁷

Issues related with the indexation of grant numbers

The availability of grant number information becomes very helpful to identify and resolve some of the issues detailed above. Grant numbers provide a far more structured and patterned source of data, although when they appear as serial number only they are naturally ambiguous. Moreover, though more structured, grant numbers nevertheless pose data problems of their own, particularly in the variations in which they appear (e.g. EY014801 also appears as: “NEI P30 Core Grant EY014801”, “NEI P30 EY014801”, “P30 EY014801”, “P30 EY-014801”, “P30EY014801”, “P30-EY014801”).

It is the frequent presence within the grant numbers of alphabetical characters, hyphens and other forms of punctuation which makes it possible to identify highly specific grant number patterns belonging to specific funding organizations, thus allowing the possibility to assign the acknowledgements in which these numbers are mentioned to these organizations. The US National Institutes of Health provide a good example of this. Their grant numbers clearly signify the individual institutes. For example, the strings “EY” (National Eye Institute) or “CA” (National Cancer Institute) provide a way to identify these institutes as funders even when they are not explicitly mentioned in the FA.

CWTS thesaurus

Considering all the inconsistencies mentioned above, CWTS has commenced with the creation of a thesaurus of funding organizations and sources extracted by WoS. The construction of this thesaurus follows a similar approach to the cleaning of affiliation data carried out for the Leiden Ranking (Waltman et al., 2012). Thus, a large extent of the FA data provided by WoS has been mapped to thesaurus entries, as such creating a new, cleaner and more workable data set, including funding organizations (Wellcome Trust, Deutsche Forschungsgemeinschaft); funding schemes (Cancer Center Support Grants, Horizon 2020), and organizations mentioned in the FA that are not primarily funding-oriented (e.g.

⁷ The extraction of Thermphos International (a former venture in the chemical industry sector) as funding body also provides another example of indexation on the basis of an unspecified, general “support”.

universities, research institutions, etc.). In the process, hierarchical connections between funding bodies have also been established, so that for instance the National Cancer Institute (US) has been identified as a child institute of the National Institutes of Health (US). The leading rule in cleaning up the data has been to retain as much of the information inherent in the data as possible, but up to a certain threshold.⁸ If a funding body does not reach this threshold, its acknowledgements are mapped to its parent agency (if possible). The goal here is to strike a balance between richness and usability of data.

Overall, in the CWTS thesaurus, more than 450 funding organizations, 230 funding schemes, and 6400 organizations have been identified and cleaned from the FA data indexed by WoS.

The thesaurus merges the funding body and grant number data, based on a set of rules (generally preferring grant number data in case of conflict). See the examples below:

Table 1. Thesaurus.

WoS funding body	WoS grant number	CWTS Thesaurus
Univ Michigan	013448-001	Univ Michigan - Ann Arbor
Univ Michigan	--	Univ Michigan - Ann Arbor
--	R01 CA122443	U.S. National Cancer Institute
Univ Michigan	R25 CA112383	U.S. National Cancer Institute

As this table shows, by taking data from both funding sources our Thesaurus becomes both more complete and more accurate.

DISCUSSION

The introduction of new FA data by WoS has opened up important possibilities of expanding the bibliometric toolset to the study of acknowledgements in general and funding information in particular. This study shows however that there are important aspects that need to be considered for the proper use and understanding of the FA data collected by WoS.

In the first place, the strong variation in the data collected and indexed by WoS requires the development of specific automated algorithms (e.g. Sirtes, 2013) and thesauri in order to be able to perform reliable studies on standardized lists of funding organizations.

Secondly, this study has shown the presence of important inconsistencies in the selection and indexation of FA data. The baseline of such a problem is that it is not straightforward what a “funding acknowledgement” actually is. The inconsistency between different types of support declarations and their consequent indexation shows that a more robust discussion on what is considered a “FA”, its typologies and its theoretical and conceptual operationalization, is still lacking. Therefore, the inconsistencies found in this paper are relatively unsurprising, since this lack of conceptualization has a direct effect on the operationalization of the term. This also raises the question if it should be the role of a data provider such as WoS to decide which FA to collect or not. In this regard, it would be much more helpful if WoS was to focus on collecting all acknowledgements from scientific publications (without deciding whether these constitute FAs or not), index all entities that appear in them, and simply leave to the

⁸ A funding body needs a minimum of 500 acknowledgements, currently.

bibliometricians and expert analysts the delineation and conceptualization of what they consider a FA.⁹

Thirdly, this study also highlights important inconsistencies in the indexation of the entities that are funding the research and those that are funded. The availability of grant numbers has proven to be a good instrument to clarify and correct potential mistakes. However, this opens questions on what is currently being indexed as “funders” in WoS. Again, this also calls for a more thorough and conceptual model of what is a FA and what are the actors involved in them (e.g. which is funder and which is funded).

Based on this study, several recommendations can be proposed for several stakeholders: researchers, funders, WoS and users of FA data:

1. Funders should provide the funded authors with clear funding statements containing a uniquely identifiable, explicitly mentioned, and standardized form of the funder name (including where possible grant numbers and other distinctive codes). Funders should also inform authors on how to clearly disclose the type of support (e.g. economic, travel, access to resources, etc.) they have received, allowing the possibility of better narrowing down on the types of support related to the funder.
2. WoS would rather strive towards being comprehensive in collecting all acknowledgements from scientific publications and extracting all entities mentioned therein, without making a priori decisions on what constitutes a FA. This would clearly contribute to a more consistent database of acknowledgements and acknowledged entities, while also expanding the bibliometric scope of their data by opening the possibility to study all types of acknowledgements (Cronin, McKenzie, & Stiffler, 1992).
3. Scientometric researchers and practitioners need to observe caution when working with FA WoS data, as not all “funding bodies” and acknowledgements that are collected by WoS necessarily constitute funders, and some of the FA metadata may also have some inaccuracies, omissions and deficiencies, making their use problematic. In addition, it is important to count with thesauri or standardized methodologies in order to be able to properly work with the FA data provided by TR (in a similar fashion as it was necessary for WoS affiliation data, cf. Fernández, et al. (1993)).

⁹ This would also open the possibility to studying what distinguishes “internal” from “external” funding; what represents a conflict of interest disclosure, and which specific types of support (e.g. access to materials or equipment, travel support) constitute “funding”.

REFERENCES

- Costas, R., & van Leeuwen, T. (2012). Approaching the “Reward Triangle”: General Analysis of the Presence of Funding Acknowledgments and “Peer Interactive Communication” in Scientific Publications. *Journal of the American Society for Information Science and Technology*, 63(8), 1647–1661. doi:10.1002/asi.22692
- Cronin, B., McKenzie, G., & Stiffler, M. (1992). Patterns of acknowledgment. *The Journal of Documentation*, 48(2), 107–122.
- Desrochers, N., Paul-Hus, A., Bowman, T., Costas, R., Haustein, S., Larivière, V., ... Tsou, A. (2015). Authorship, patents, citations, acknowledgments, tweets, reader counts and the multifaceted reward system of science. In *Proceedings of the American Society for Information Science and Technology*. Retrieved March 29, 2016 from <https://www.asist.org/files/meetings/am15/proceedings/openpage15.html>
- Díaz-faes, A. A., & Bordons, M. (2014). Acknowledgments in Scientific Publications □: Presence in Spanish Science and Text Patterns Across Disciplines. *Journal of the Association for Information Science and Technology*, 65(9), 1834–1849. doi:10.1002/asi.23081
- Fernández, M. T. T., Cabrero, A., Zulueta, M. A. A., & Gómez, I. (1993). Constructing a relational database for bibliometric analysis. *Research Evaluation*, 3(1), 55–62.
- Honk, J. van, Calero-Medina, C., & Costas, R. (2016). A case study on the accuracy of Web of Science funding acknowledgement data: PLoS. In preparation.
- Paul-Hus, A., Desrochers, N.; & Costas, R. (2016). Characterization and description of funding acknowledgement data in Web of Science: key issues in their consideration as a research tool. Submitted to *Scientometrics*.
- Rigby, J. (2011). Systematic grant and funding body acknowledgement data for publications: new dimensions and new controversies for research policy and evaluation. *Research Evaluation*, 20(5), 365–375.
- Sirtes, D. (2013). Funding Acknowledgments for the German Research Foundation (DFG). The dirty data of the Web of Science database and how to clean it up. In J. Gorraiz, E. Schiebel, C. Gumpenberger, M. Hörlesberger, & H. Moed (Eds.), *Proceedings of the 14th International Society of Scientometrics and Informetrics Conference* (pp. 784–795). Vienna: AIT GmbH.
- Thomson Reuters. (2015). Thomson Reuters Bibliographic Policy Funding Acknowledgements. Internal document.
- Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E. C. M., Tijssen, R. J. W., Eck, N. J. Van, ... Wouters, P. (2012). The Leiden Ranking 2011 / 2012 □: Data Collection, Indicators, and Interpretation. *Journal of the American Society for Information Science and Technology*, 63(12), 2419–2432. doi:10.1002/asi.22708
- Wang, J., & Shapira, P. (2011). Funding acknowledgement analysis: an enhanced tool to investigate research sponsorship impacts: the case of nanotechnology. *Scientometrics*, 87(3), 563–586.



Open Data in Global Environmental Research: Findings from the Community¹

Birgit Schmidt*, Birgit Gemeinholzer and Andrew Treolar

*bschmidt@sub.uni-goettingen.de

State and University Library, University of Göttingen, Platz der Göttinger Sieben, Göttingen, 37070 (Germany)

INTRODUCTION

This paper presents findings from the Belmont Forum's survey on Open Data which targeted the global environmental research and data infrastructure community (Schmidt, Gemeinholzer & Treolar, 2016). It highlights users' perceptions of the term "open data", expectations of infrastructure functionalities, and barriers and enablers for the sharing of data. A wide range of good practice examples was pointed out by the respondents which demonstrates a substantial uptake of data sharing through e-infrastructures and a further need for enhancement and consolidation. Among all policy responses, funder policies seem to be the most important motivator. This supports the conclusion that stronger mandates will strengthen the case for data sharing.

The Belmont Forum, a group of high-level representatives from major funding agencies across the globe, coordinates funding for collaborative research to address the challenges and opportunities of global environmental change. In particular, the E-Infrastructure and Data Management Collaborative Research Action has brought together domain scientists, computer and information scientists, legal scholars, social scientists, and other experts from more than 14 countries to establish recommendations on how the Belmont Forum can implement a more coordinated, holistic, and sustainable approach to the funding and support of global environmental change research.

METHODS

In the context of the working group on Open Data researchers from various science communities, interested laypersons, government employees, and others were invited to share their views and experiences on data publishing, access and (re)use. The main aim of the survey was to learn more about:

- Key open data activities in various communities dealing with global environmental change to identify leading examples of best practice from a user perspective;
- Areas where users' desire to share could be enhanced by new/other developments;
- Barriers to "open data sharing" from a user perspective (as either a data provider or data user).

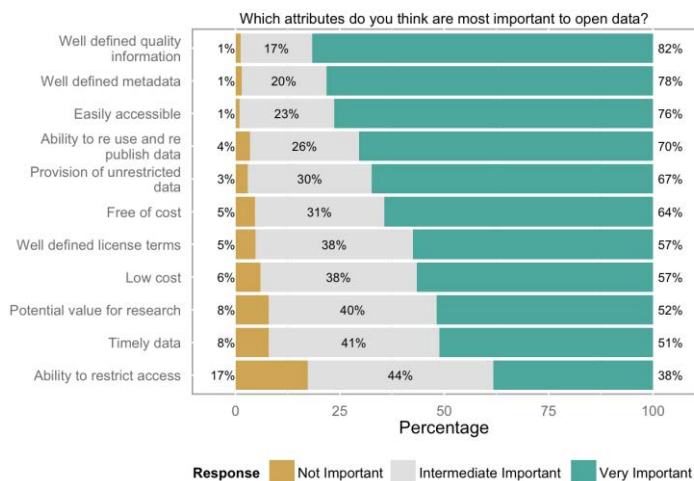
From September to November 2014, the survey collected over 1,300 responses based on the distribution of the survey to about 20 disciplinary and professional mailing lists, and to all the authors of a well-renowned open access publisher, central to the research area. All of the 19 questions of the survey were non-mandatory. For the analysis the statistics software R and in particular the Likert package were used. All data are available via the Zenodo repository (Schmidt et al, 2016).

¹ This work was supported by the Belmont Forum's E-Infrastructure and Data Management Research Action.

KEY FINDINGS

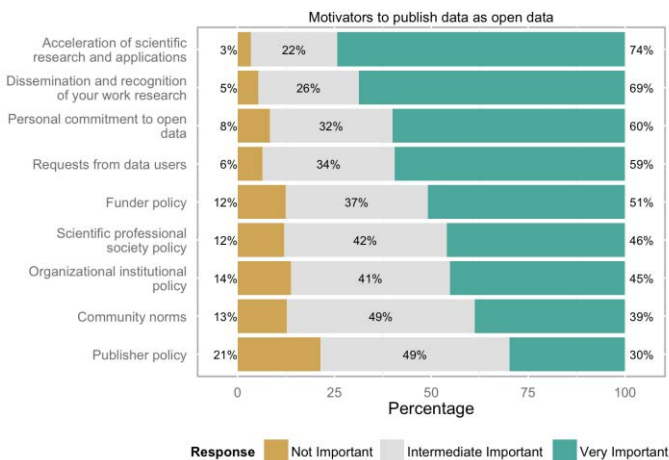
Instead of providing a definition of “open data” the survey assessed the user perception associated with the term (compare Fig. 1). The answers highlight the importance of information which enables the user to assess the quality of data, to select data based on metadata, and to easily access and reuse the data. The ability to restrict access was lowest in the ranking of desirable attributes, which fits the intuitive idea of openness. However, nearly 2/5 off all respondents still considered the option to restrict data as a very important attribute.

Figure 1: Perceived properties of open data.



Motivators and barriers to publish data as open data were studied in the survey. The main desires to publish data as open data arose from research-intrinsic motives ranging from general considerations, i.e. the acceleration of scientific research and applications, to personal motivations, i.e. dissemination and recognition of research results, personal commitment to open data and requests from data users (cf. Fig. 2). Among the three types of data professionals which responded to the survey (data user, data provider, data manager) data managers’ personal commitment to open data seem to be significantly higher.

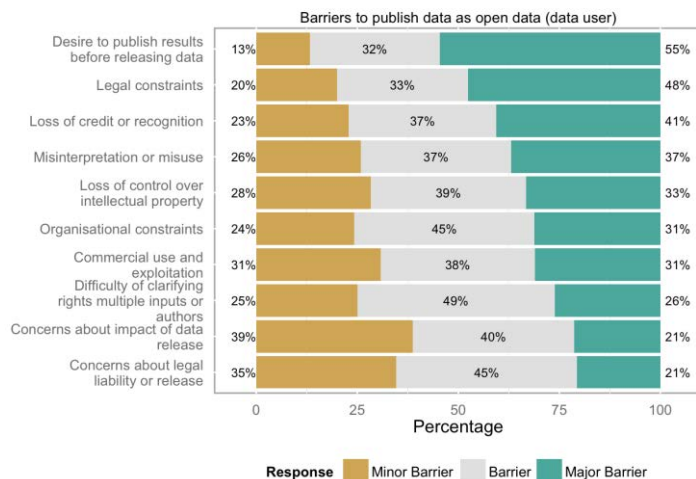
Figure 2: Motivators to publish data as open data.



Overall, the most important barrier for publishing data as open data were the desire to publish results before releasing data, legal constraints, loss of credit or recognition and possible

misinterpretation or misuse. Concerns about legal liability for data or release of data were least pronounced. In addition, the desire to publish results before releasing data was somewhat more prevalent at early stages of a research career.

Figure 3: Barriers to publish data as open data.



In addition, the survey explored where the community accesses and/or publishes data, and a wide range of good practice examples was pointed out by the respondents (several of these data repositories are currently added to the re3data.org registry) which demonstrate a substantial uptake of data sharing and reuse through data e-infrastructures in the global environmental change community. A need for further enhancement and consolidation can be derived from the respondents' expectations about functionalities of infrastructures and desires expressed about access to specific types of data.

CONCLUSIONS

Based on the findings of the survey the following actions were recommended to the Belmont Forum:

- Funders should make open data archiving mandatory, while taking into account the main motivators revealed by the survey.
- Scientific merit as well as accelerating research and applications are still the main motivators for publishing data; thus ethics of data sharing and reuse should be taken into account when proposing guidelines for open data sharing and re-use.
- Support and training activities should be supported in concerted ways, targeting researchers as well as current and future data and information professionals.
- Interoperability between infrastructures should be further facilitated, taking into account generic requirements (e.g. providing links to publications and funder information) as well as disciplinary norms and standards (e.g. vocabularies, metadata standards).

REFERENCES

Schmidt, B., Gemeinholzer, B. & Treloar, A. (2016). Open Data in Global Environmental Research: The Belmont Forum's Open Data Survey. *PLoS ONE*, 11(1), e0146695. doi:10.1371/journal.pone.0146695.

Schmidt, B., Gemeinholzer, B., Treolar, A., Hodge, J., Santanchè, A. & Oakley, K. (2015). Belmont Forum Open Data Survey. Zenodo, doi:10.5281/zenodo.16384



Examining data access and use in science¹

Erjia Yan^{*}, Mengnan Zhao^{*}

^{*}ey86@drexel.edu; mz438@drexel.edu

College of Computing and Informatics, Drexel University, 3141 Chestnut Street, Philadelphia, PA (U.S.A.)

ABSTRACT

In this research-in-progress paper, we provide preliminary evidence of data access and use in scientific literature based on a content analysis of 600 stratified sampled PLOS ONE publications. Results show that data access and use varied greatly from one paper to another in terms of how datasets were collected, referenced, and curated.

INTRODUCTION

Data have progressively become an integral component in modern science—thanks to the increased permeability of disciplinary boundaries, the enhanced human mobility, and the advanced technologies to process, analyze, and curate large scientific data. Scientists can now form interdisciplinary, international collaborative teams surrounded by data to conduct the so called data-centric or data intensive research (Tansley & Tolle, 2009).

In science, there is a growing awareness of data access and sharing. As early as 2004, scholars have advocated for an international framework to promote data accessibility (Arzberger et al., 2004). It is argued that data sharing helps develop a democratic society (Harrison et al, 2012), enhances the transparency of scientific research particular for those sensitive topics such as climate change (Parmesan & Yohe, 2003), allows for reproducing and validating research (Bradley, 2009; Nosek et al., 2015), and unleashes the potential of data to solve complex societal issues such as diabetes (Zimmet et al., 2001). Realizing these benefits, a number of scientific journals and funding agencies have begun mandating making data freely available to the public: for instance, *Nature* requires authors to “make materials, data, code, and associated protocols promptly available to readers without undue qualifications” (Nature Editor, n.d.), and likewise the National Science Foundation of the U.S. expects investigators to “share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants” (National Science Foundation, n.d.). Organizations have also made an effort of indexing data such as Thomson Reuters’ *Data Citation Index* (Thomson Reuters, n.d.) or SageCite by University of Bath, U.K. (Lyon, 2010).

Despite these efforts, access to data is still highly inconsistent and even obscure. Data can be formally curated in journal-specific digital repositories or institutional archives that are typically assigned with DOIs or URLs, or informally stored in personal computers and

¹ This work was made possible in part by the Institute of Museum and Library Services (Grant Award Number: RE-07-15-0060-15), for the project titled “Building an entity-based research framework to enhance digital services on knowledge discovery and delivery”. In addition, this work was also supported by the National Consortium for Data Science (NCDS) for the Data Fellows program.

This work is licensed under a Creative Commons License: Attribution-NonCommercial-NoDerivatives 4.0 International

servers. As a result, data are referenced in unsystematically ways in scientific literature: they can be formally cited, or simply be mentioned in paragraphs, footnotes, endnotes, and acknowledgements. A 2014 study on oceanography data access found that data are more likely to be mentioned in the text than been formally cited (Belter, 2014). Thus, merely using citation indices is insufficient to capture the different ways of data access and use. Instead, full-text publications provide the crucial context for this purpose. This research-in-progress paper reports a preliminary set of results on several key aspects of data access and use in science.

METHODS

DATA SOURCE

The data set used in the study contained open access, full-text papers from PLOS ONE. The access to the data set is provided by the PubMed Central Open Access Subset (<http://www.ncbi.nlm.nih.gov/pmc/tools/openflist/>) and is publically available. We sampled 50 papers randomly from 12 defined discipline for papers published during 2014-2015, which resulted in 600 papers in total.

CODING PROCEDURES

We employed content analysis as the research instrument because it is an effective method to discover quantitative patterns from textual corpora (Bauer, 2000; Herring, 2010; Krippendorff, 2012). In content analysis, coding is the crucial link between data collection and data interpretation, allowing researchers to use a set of guidelines (i.e., coding schemes) to systematically make sense of data.

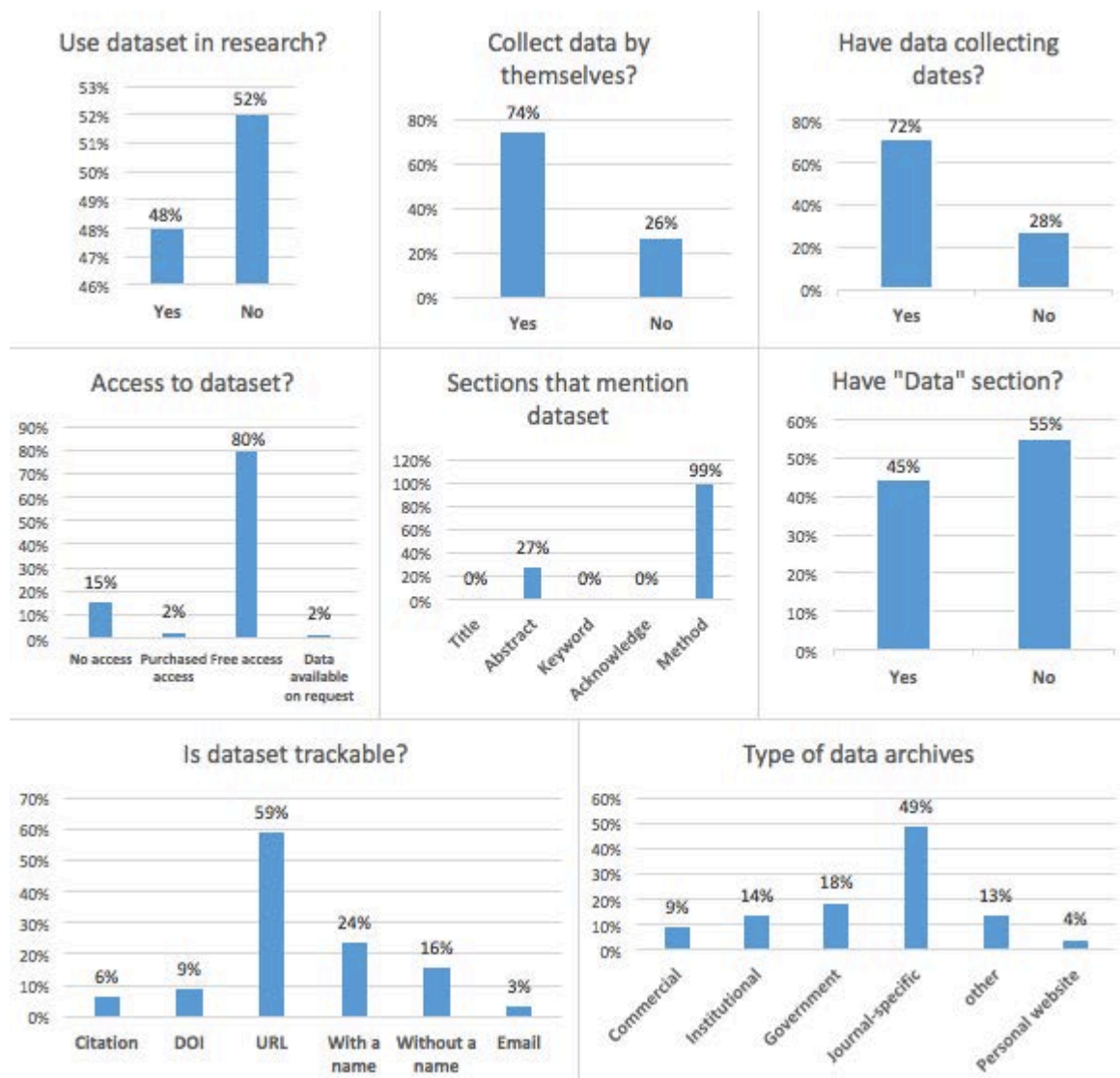
We first created a draft coding scheme and then adopted the grounded theory approach and applied the draft scheme to a subset of the data with the goal to identify previously unnoticed yet valuable patterns—this process helped us complement the coding scheme and the finalized version is shown in Table 2. Most of the coding items are pre-coordinated, while new codes may emerge during coding, which are referred to as emergent codes. We marked the emergent codes with “*” and kept refining the coding scheme during the whole coding process. The final coding scheme is shown in Figure 1. It should be noted that while most PLOS ONE publications may have used data to certain extent (i.e., quantitative research), in this study, we operationalized data as datasets—data that were stored in certain formats or media (for instance, a paper that used a statistical analysis without explicitly mentioning datasets is therefore not considered as a paper that used data in research). Because of the unambiguous and self-explanatory nature of the coding scheme as well as an obtained inter-rater reliability of 1 on a small sample of the data, one coder coded all 600 papers.

Figure 1. Coding Scheme

1. Use data in research	4. Data access	6. Section of data mention
1.1 Y	4.1 No access	6.1 Title
1.2 N	4.2 Purchased access	6.2 Abstract
2. Data section	4.3 Free access	6.3 Keyword
2.1 Y	*4.4 Data available on request	6.4 Acknowledge
2.2 N	5. Data tracking	6.5 Method
3. Data collection	5.1 Citation	7. Type of archives
3.1 Collecting data on their own	5.2 DOI	7.1 Commercial
3.1.1 Data collection date	5.3 URL	7.2 Institutional
3.1.1.1 Y	5.4 With a name	7.3 Governmental
3.1.1.2 N	5.5 Without a name	7.4 Journal-specific
3.2 Using public data set	*5.6 Email	7.5 Other
		*7.6 Personal website

RESULTS

After finishing the coding work, we went through the coded articles and counted the number of articles in each coding category, with the results shown in Figure 2.

Figure 2. Percentage of articles in each coding category

Among the 600 articles in the data set, 52% (or 312 articles) used datasets in their research. Within the 312 articles that used datasets, only about half of them had data or data related sections; the others just mentioned data sources, data collection methods, or data descriptions in method sections. Within the 312 articles, 74% collected data and created datasets by themselves; among these, 72% provided the date or the time period when data were collected. The numbers demonstrate that a majority of studies are inclined to create datasets and use their own data, rather than reuse previously created and curated data from others' research.

For data attribution methods, citations and data identifiers are considered the most normative by facilitating ways of data tracking; however, only 6% and 9% of the articles respectively attributed data in such formal ways. Meanwhile, 60% of the articles provided URL to locate datasets. Most of the provided URLs worked at the present time, but the concern is that once an URL expires, we will lose track of the datasets. Furthermore, 24% articles just provided

This work is licensed under a Creative Commons License: Attribution-NonCommercial-NoDerivatives 4.0 International

the names of the datasets, some of which are quite unique to be located while others may refer to multiple data entities. Some articles also attached an email address with the dataset so that readers can send email requests for data.

In regards to means of data storage, nearly half of them saved data in the journal website as attachments to papers, followed by housing data in governmental (18%), institutional (14%) and commercial (9%) repositories. In addition, 4% articles hosted data in researchers' personal websites.

CONCLUSIONS

In this research-in-progress paper, we provided preliminary evidence of data access and use in scientific literature based on a content analysis of 600 stratified sampled PLOS ONE publications. Results showed that data access and use varied greatly from one paper to another in terms of how datasets were collected, referenced, and curated. The next step in this research project will involve the identification of disciplinary characteristic of data access and use as well as the design of inclusive indicators to comprehensively capture the full-spectrum of data impact.

REFERENCES

- Arzberger, P., Schroeder, P., Beaulieu, A., & Bowker, G. (2004). An international framework to promote access to data. *Science*, 303(5665), 1777.
- Bauer, M. (2000). Classical content analysis: A review. In M. Bauer & G. Gaskell (eds.), *Qualitative Researching with Text, Image and Sound* (pp. 131-151). Thousand Oaks, CA: Sage.
- Belter, C. W. (2014). Measuring the value of research data: a citation analysis of oceanographic data sets. *PloS one*, 9(3), e92590.
- Bradley, J. C. (2007). Open notebook science using blogs and wikis. Retrieved June 1, 2016 from <http://precedings.nature.com/documents/39/version/1>
- Harrison, T. M., Guerrero, S., Burke, G. B., Cook, M., Cresswell, A., Helbig, N., ... & Pardo, T. (2012). Open government and e-government: Democratic challenges from a public value perspective. *Information Polity*, 17(2), 83-97.
- Herring, S. C. (2010). Web content analysis: Expanding the paradigm. In J. Hunsinger, M. Allen, & L. Klastrup (Eds.), *The International Handbook of Internet Research* (pp. 233-249). Berlin: Springer Verlag.
- Krippendorff, K. (2012). *Content analysis: An introduction to its methodology*. Sage.
- Lyon, L. (2010). UK Digital Curation Centre: enabling research data management at the coalface. Microsoft Research Faculty Summit 2010, Redmond, Washington, USA, July 2010.
- National Science Foundation (n.d.). Dissemination and Sharing of Research Results. Retrieved March 10, 2016 from <https://www.nsf.gov/bfa/dias/policy/dmp.jsp>

Nature editor (n.d.). Availability of data, material and methods. Retrieved March 10, 2016 from <http://www.nature.com/authors/policies/availability.html>

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., ... & Contestabile, M. (2015). Promoting an open research culture: Author guidelines for journals could help to promote transparency, openness, and reproducibility. *Science* (New York, NY), 348(6242), 1422.

Parnesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421(6918), 37-42.

Tansley, S., & Tolle, K. M. (Eds.). (2009). *The fourth paradigm: data-intensive scientific discovery* (Vol. 1). Redmond, WA: Microsoft research.

Thomson Reuters (n.d.). Data Citation Index. Retrieved March 10, 2016 from http://wokinfo.com/products_tools/multidisciplinary/dci/

Zimmet, P., Alberti, K. G. M. M., & Shaw, J. (2001). Global and societal implications of the diabetes epidemic. *Nature*, 414(6865), 782-787.

SMS: a linked open data infrastructure for science and innovation studies^{*}

Peter van den Besselaar^{*}, Ali Khalili^{**}, Al Idrissou^{*}, Antonis Loizou^{**},
Stefan Schlobach^{**}, Frank van Harmelen^{**}

p.a.a.vanden.besselaar@vu.nl

^{*} Network Institute & Department of Organization Sciences
Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

^{**} Network Institute & Department of Computer Science
Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

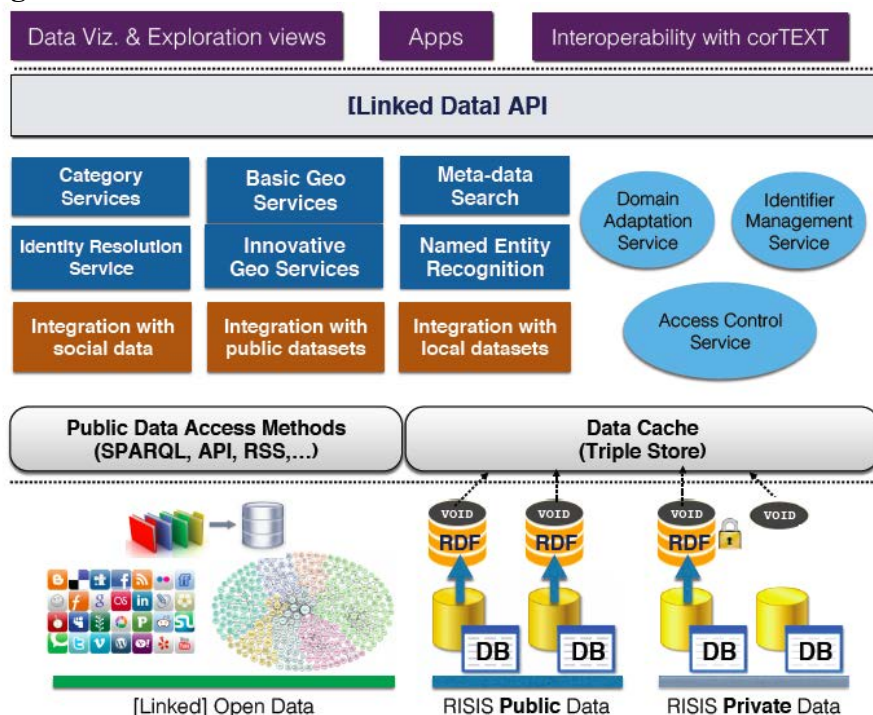
ABSTRACT

In this paper we describe a data integration infrastructure for Science Technology and Innovation (STI) studies developed within the context of the RISIS project. We outline its architecture and functionalities. In the full paper, we will show the use of the infrastructure in a complex research project. At the conference we will give a demonstration.

INTRODUCTION

In this paper we describe a data integration platform for science technology and innovation data using semantic web technology and focusing on (but not restricted to) linked open data (Beek et al 2016). Figure 1 shows the basic architecture.

Figure 1: SMS architecture



^{*} EC grant 313082, the RISIS project: Research Infrastructure for Science and Innovation Studies.

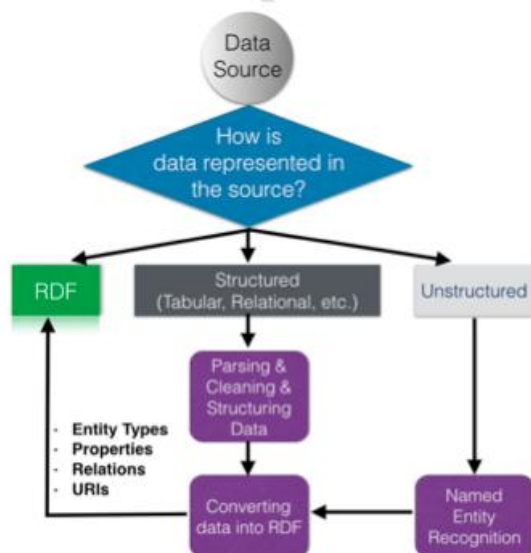
Why is such infrastructure needed? Up to now, STI studies are either *rich* but small scale (qualitative case studies) or large scale and *under-complex* – because they generally use only a single dataset like Patstat, Scopus, WoS, OECD STI indicators, etc., and therefore deploying only a few variables – determined by the data available. However, progress in the STI research field depends in our view on the ability to do large-scale studies with often many variables specified by relevant theories: There is a need for studies which are at the same time big *and* rich. To enable that, combining and integration of STI data and beyond is needed – in order to exploit the many data that are „out there“ in an innovative and meaningful way. That is why the core of the infrastructure is the conversion of different datasets in the same open format: from tabular data, text data and web data to RDF (Beek et al 2016).

This emphasis on data integration is also visible in other research fields. That enables to build a data infrastructure partly by reusing existing tools. Within the RISIS project we develop the *SMS platform for data integration and data enrichment* by combing those existing tools with specific tools newly developed for the STI field. The SMS platform partly implemented now; we aim at providing a complete beta version later on this year, as part of the RISIS S&I data infrastructure (www.risis.eu). The following functions can be distinguished in the SMS platform:

Pre-processing

Pre-processing data and converting data into the RDF standard for linked open data (Figure 2). For example, PDF files can be converted into TXT, and through *Named Entity Recognition* relevant entities like people, organizations, countries, etc. are identified. Additional text processing (e.g., term extraction) may identify attributes. A concrete example is recognizing research institutions and universities in a researcher’s CV, using name recognition by linking the CV to databases with background knowledge such as DBpedia. The resulting data are then converted into RDF. Structured data (e.g., Excel files) are parsed and cleaned. And then converted into RDF.

Figure 2: From heterogeneous data to RDF

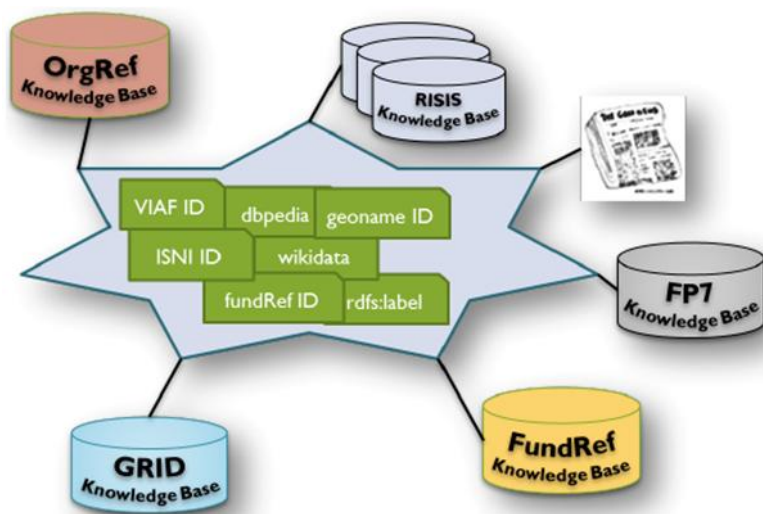


Linking data

The next step is *linking* the data. If entity *identifiers* are available, the linking is easy. If not, a variety of techniques can be used, from (fuzzy) *string matching* to *deploying attributes*

available in the different databases. If names occur in different languages, resources like DBpedia can be used to match. If two entities have different names, but similar characteristics, they may be in fact the same entity. However, whether entities are considered the same, depends on the perspective: sometimes two organizations (e.g. departments) can be the same – because they are parts of the same organization (university). But if one wants to compare departments, this is not the case. We are currently experimenting with a series of datasets on research organizations, in order to compile basic reference sets of research organizations. This is done through interlinking different datasets through knowledge resources on the web (Figure 3).

Figure 3: Linking data through web knowledge resources



There is also the problem of disambiguation of person names, which is (in our field) mainly related to publications and patents, and for which specific tools are being developed (Sandström & Sandström 2009; Gurney et al 2012). One of the questions addressed is how complete disambiguation needs to be given the questions posed in a research project (Van den Besselaar & Sandström, forthcoming). Within the SMS platform we do not tackle this problem field for the moment, and the researcher can use existing tools – together with still quite some manual work.

Geo-services

An interesting possibility is linking through geo-location: if two entities have the same geo-location, they may be related (or identical). Geo-locating has an additional advantage, as it is also an instrument to enrich data: many other (open) data provide variables measures at some level of geographical aggregation: e.g., environmental data, educational data, or socio-economic data. Therefore the platform provides a variety of *geo-services*.

We illustrate this with an example of a service to determine the geographical location if one knows an address (or even only a name). The system is based on a series of open geo-resources, such as GeoNames and OpenStreetMap (figure 4). In the top left part of the screen the address “Vrije Universiteit Amsterdam” is inserted, and the service has as output various maps and, in the bottom right, the geo-characterization of the inserted address at eleven levels. Figure 4 shows the various administrative boundaries where Level 8 represents LAU 2.

By integrating these resources, the service can give for an entity's address the geo-location at 11 different levels, which then can be used to link the entity to other (often statistical) data. Of course the platform can be used to do this for larger amounts of addresses, and the output then is not on the screen, but in a tabular form. In the future we aim at adding different distance concepts, such as travel distance (time, frequency, price, etc.).

Figure 4: Geo-locating services

Category services

As datasets may use different category systems for the attributes, linking data requires a mapping of these category systems or „vocabularies“ (Figure 5). A good example are the different systems that are used for classifying research fields, e.g., in the Web of Science and in OECD R&D statistics. A *category service* would enable the data user to select which one

The screenshot displays a web interface for geo-locating services. It features several panels:

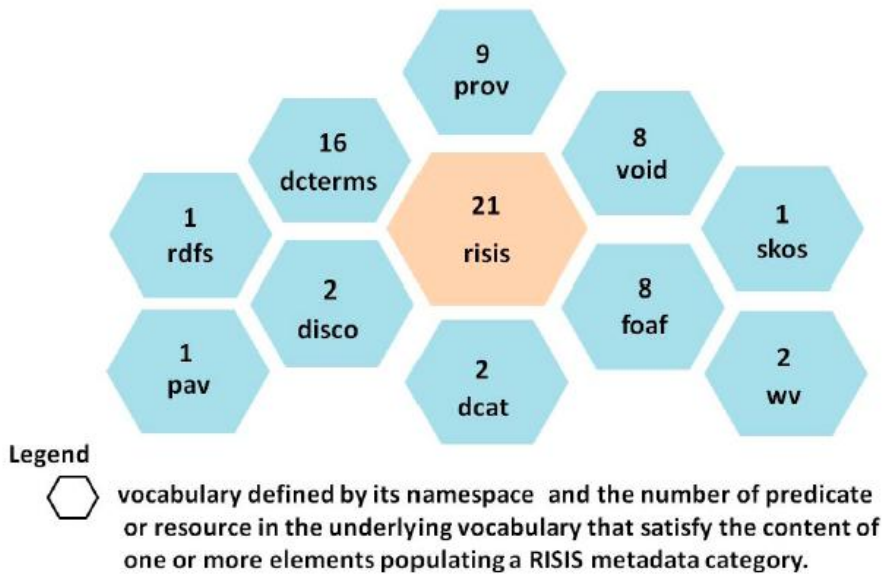
- Address to Coordinates:** A text input field contains "vrije universiteit amsterdam". Below it, the "Formatted Address" is "Aula, De Boelelaan 1105, 1081 HV Amsterdam, Netherlands". The "Latitude" is 52.3344753 and the "Longitude" is 4.8658302.
- Coordinates to Municipality:** A map shows the location in Amsterdam, with a red pin and a purple polygon indicating the municipality boundary. The code "NL270363" is displayed.
- NUTS to Municipality:** A table lists 27 items found, with columns for "Name" and "Code".

Name	Code
Almere	NL240034
Abcoude	NL260305
De Ronde Venen	NL260736
Amstelveen	NL270362
- Municipality to FUA:** A table shows the mapping for Amsterdam.

Name (FUA)	Code (FUA)	Is Core ?
Amsterdam (Amsterdam)	NL270363 (NL002)	✓
- Coordinates to NUTS:** A map of the Netherlands shows the location of Amsterdam, with a red pin and a purple polygon. The code "NL326" is displayed.
- PointToOSMAdminBoundaries:** A list of 11 administrative boundaries is shown, including "Nederland", "Noord-Holland", "Stadsregio Amsterdam", "Metropoolregio Amsterdam", "Amsterdam -> LAU-2", "Stadsdeel Zuid", "Amsterdam", and "Zuideramstel".

classification he/she wants to use. And the system would then do the mapping between the different classifications. For this, we deploy existing vocabularies available on the web. One can also think of other classification schemes that can be mapped, e.g., of professions, of jobs, of types of organizations, and so on. As many developments are taking place, the SMS platform may use what is available. E.g., within the RISIS project work is done on classifications of companies, and of research organizations. The RISIS metadata system will be of help here.

Figure 5: Integrating vocabularies



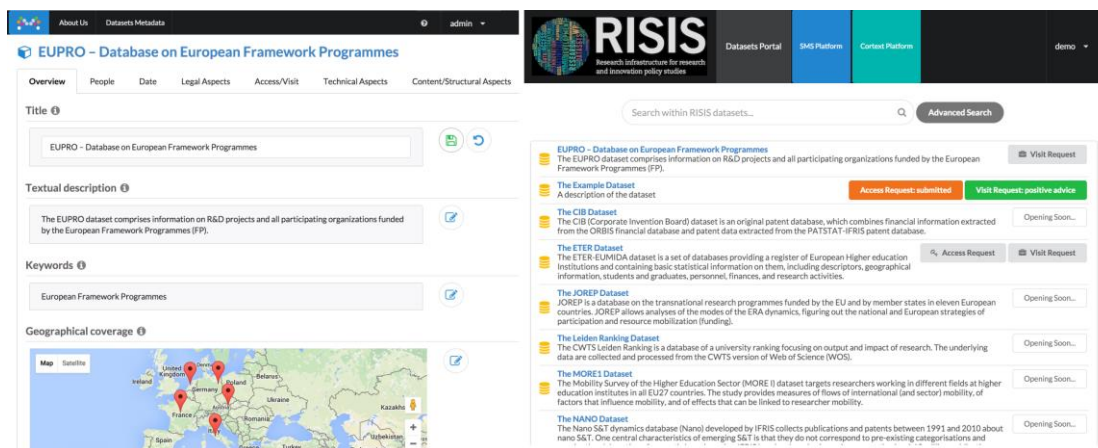
Improving quality and data enrichment

Linking can also be used to *improve quality* of the data and *enrich* them. Linking the two sets may increase the number of variables, but also may reveal discrepancies in variable values, and the user should then be able to decide what the more reliable source is. Quality improvement follows from detecting value differences or similarities between datasets. *Quality assessment* using among other *provenance* will be implemented too: What was done with the data, and how. This should be transparent for the user.

Metadata

The platform offers a metadata system, which is also linked to open data in order to have advanced search facilities. The metadata system is also a tool to support data integration, due to the fact that the dataset owner is stimulated to use URLs in the metadata (figure 6). And it is supported by the category services discussed above. (For more details: Idrissou et al (2015).

Figure 6: The RISIS/SMS metadata system



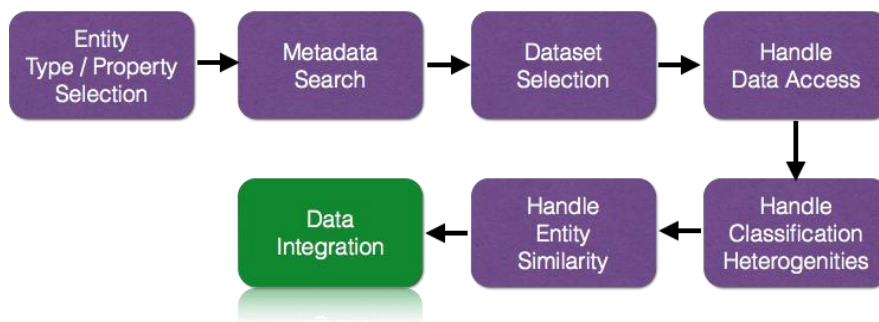
Access control

The platform provides access to a variety of datasets, of which some are open, some are proprietary and require e.g., subscription, and other are confidential. As data are only partly open, access control is provided and essential. With the type of data we are focussing on, privacy issues and legal issues may easily come up.

The workflow

From the users' perspective, the platform does two things. Firstly there is the workflow to identify data needed by the user to do a research project. This goes from identifying the entities and the variables (properties) needed. Through the metadata search the relevant datasets can be selected. If access can be given, steps follow like classification matching and disambiguation, and then the data can be integrated. The workflow is represented in figure 7.

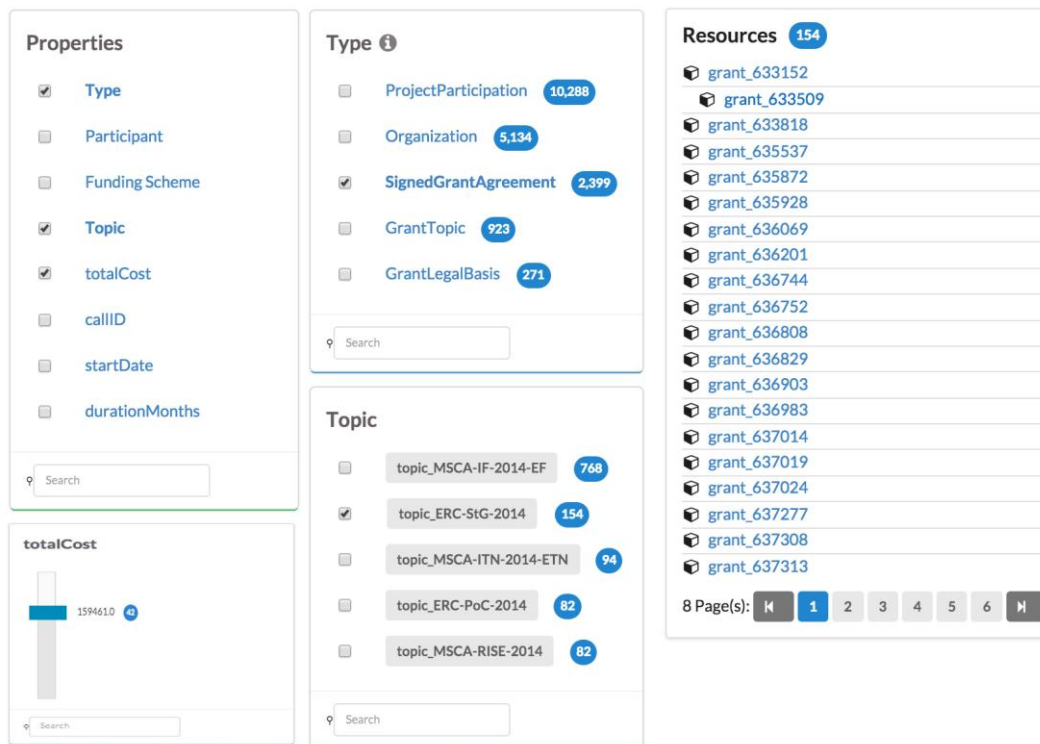
Figure 7: The users workflow



Secondly, when the integrated data are available, the user wants to have a dataset to do the analysis and visualization. (Standard) queries are provided to get the required data into the required format. This sounds simpler than it is, but experience with other data integration platforms show that the user needs support by specialists to query the platform. This suggests that it is indeed more an infrastructure than a tool. The output can have various formats, to enable deployment of general or specific analytical tools. A specific interface will be developed to connect the SMS platform to the Cortext platform (www.cortext.fr).

Data can also be browsed, in order to get a more qualitative feeling for the data. The *facet browser* is used for this (figure 8). Faceted browsing is particularly useful when you would like to present users with multiple entry points into a dataset or when there is no expectation that they know what they are looking for beforehand. It allows users to explore the space of potential items by choosing the refinements in any order.

Another use of the facet browser is when searching for information for more qualitative studies. The linked nature of the data enable to search for rich information about the entities one is interested in.

Figure 8. An example of SMS faceted browser

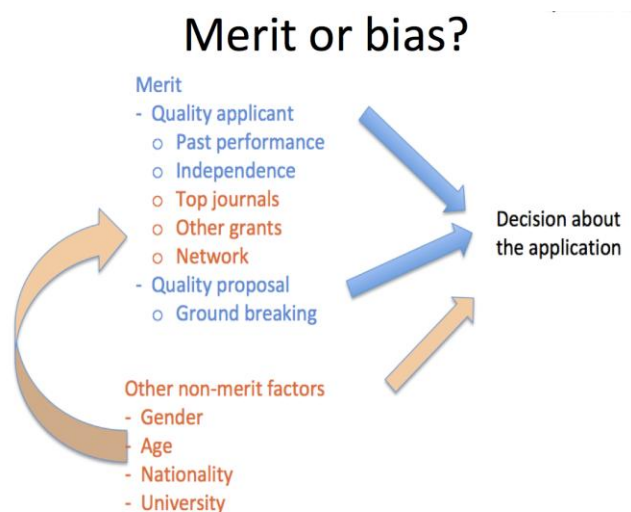
THE DEMONSTRATOR

Many parts of the platform are already implemented and tested. Currently we are finalizing the beta-version of the platform, and the planning is that after this summer, the platform is available for the first users at <http://sms.risis.eu>.

We show the use of SMS in a demonstrator project, investigating gender bias in grant allocation. In this project we try to find out whether gender of applicants influences the grant decision. In order to answer that question, one needs to bring in a multitude of variables that may influence the decision – apart from gender. This may be variables representing merit – such as measures of scholarly performance, but also variable that measure performance in a possibly gendered way, such as the collaboration network. And it needs to include personal characteristics that can influence the decision, such as age, nationality, and so on. The model we use (figure 9) includes quite some – theory driven – variables. These variables come from a variety of data sources:

- From Web of Science: Bibliometric performance scores
- Quality of the applicants network: Organizations mentioned in the CV (PDF), and ranking of those organizations from Leiden Ranking (Excel)
- Earlier grants: from CV (PDF)
- Host institution from admin file (Excel) and ranking of host institution from administrative file (Excel)
- Personal characteristics from admin file (Excel)
- Linguistic categories in evaluation: Term extraction from review forms (PDF)

We used the SMS platform for pre-processing, for converting into RDF, for entity recognition and linking. The output is a data file for analysis.

Figure 9: the GendERC project – analytical model

Preliminary findings suggest that gender bias indeed exists, but different in the different disciplines. But for this paper, it is more interesting how it was done then what comes out (Van den Besselaar et al 2016; Van den Besselaar 2016)

CONCLUSIONS

We expect that platforms like SMS will enable research within the STI field that was not possible before. Studies can become large-scale, can including many more variables than traditionally has been the case. More and more appropriate data can be exploited. Within the (life) sciences, instrumentalities and infrastructures have radically changed the way research is done (de Solla Price 1984). In the social sciences and humanities this has been much less that case; but that may change in the near future. New data integration and enriching infrastructures may open the space of new forms of social science. As Nicholas Christakis (2013) wrote: “Let’s shake up the social sciences”.

REFERENCES

- Beek W, Rietveld L, Schlobach S, Van Harmelen F, LOD Laundromat; Why the semantic web needs centralization (even if we don’t like it). *IEEE Internet Computing*, March-April 2016
- Christakis N, Let’s shake up the social sciences” *New York Times* (July 19, 2013),
- Ciccarese P, S. Soiland-Reyes, K. Belhajjame, A. J. Gray, C. Goble, and T. Clark. Pav ontology: provenance, authoring and versioning. *Journal of biomedical semantics*, 4 (2013) 1,1-22,
- Daraio C, M. Lenzerini, C. Leporelli, H. F. Moed, P. Naggar, A. Bonaccorsi, and A. Bartolucci. Data integration for research and innovation policy: an ontology-based data management approach. *Scientometrics*, 1-15, 2015.
- de Solla Price D, The science/technology relationship, the craft of experimental science and policy for the improvement of high technology innovation. *Research Policy* **13** (1984) 3-20
- Groth, P, A. Loizou, A. J. Gray, C. Goble, L. Harland, and S. Pettifer. Api-centric linked data integration: The open fPHACTSg discovery platform case study. *Web Semantics: Science, Services and Agents on the World Wide Web* **29** (2014) 12-18

- Gurney T, Horlings E, van den Besselaar P, Author Disambiguation Using Multi-Aspect Similarity Indicators. *Scientometrics* **91** (2012) 435-449
- Idrissou, AK, Khalili A, Hoekstra R, van den Besselaar P, Managing metadata relevant for research and innovation studies: The RISIS case. Paper *Whise workshop* 2016
- Khalili A, A. Loizou, and F. van Harmelen. Adaptive linked data-driven web components: Building flexible and reusable semantic web interfaces. *Extended Semantic Web Conference (ESWC)* 2016.
- Sandström, U., & Sandström, E., The field factor: Towards a metric for academic institutions. *Research Evaluation* **18** (2009) 243–250.
- Van den Besselaar, P, *Predicting panel scores by linguistic analysis*. Paper submitted to the STI conference, Valencia 2016
- Van den Besselaar & Sandström, *What is the required level of data cleaning? A research evaluation case*. forthcoming
- Van den Besselaar P, Schiffbanker H, Snaström U, Holzinger F, Polo L, *Explaining gender bias in grant selection – the ERC starting grants case*. Paper submitted to the conference on Gender and Higher Education, Paris 2016



Data Citation Policies of Data Providers within the scope of Longitudinal Studies in Life Course Research¹

Anke Reinhold^{*}, Marc Rittberger^{**} and Nadine Mahrholz^{***}

^{*} reinhold@dipf.de

German Institute for International Educational Research (DIPF), Schloßstraße 29, Frankfurt, 60486 (Germany)

^{**} rittberger@dipf.de

German Institute for International Educational Research (DIPF), Schloßstraße 29, Frankfurt, 60486 (Germany)

^{***} nadine.mahrholz@uni-hildesheim.de

University of Hildesheim, Universitätsplatz 1, Hildesheim, 31141 (Germany)

ABSTRACT

In this article, a small-scale case study analyzing the nature of data citation policies within the scope of longitudinal studies in life course research is presented. The sample consists of eight data providers from Europe, North-America and Australia and was evaluated with regard to eight criteria which potentially affect data citation behavior of researchers in the field, for example the wording of data citation obligations or sanctions for not citing research data in accordance to given requirements. The study demonstrates that research data providers follow a wide range of approaches to data citation, especially in terms of data citation location within a publication as well as disposal obligations for data-related publications. However, this diversity might lead to inconsistency in data citation behaviour and also to a general lack of comparability of data citation quantity and quality as relevant factors in research evaluation.

INTRODUCTION

In order to meet the requirements of funding organisations or policy makers, the scientific output of researchers, research groups, institutions and even countries is regularly tracked by indicators that measure for example citation rates or citation impact. With the rise of altmetrics, attention in research monitoring has also shifted towards research activities that are – exclusively or complementarily – visible on the social web. However, citation analysis is still mainly focused on publication-related research output and so far only a few works have discussed the distinctiveness of research data as a considerable factor in citation analysis and research evaluation. For example, quantitative analyses of the Data Citation Index (DCI) (Thomas Reuters) (e.g. Peters, Kraker, Lex, Gumpenberger & Gorraiz, 2015; Robinson-García, Jiménez-Contreras & Torres-Salinas, 2015) as well as subject-specific publication depositories (Mooney, 2011; Mooney & Newton, 2012) have shown a general uncitedness of research data in the social sciences and the humanities, despite the fact that sharing research data can be associated with higher citation rates (Piwowar, Day & Fridsma, 2007).

Studies analysing the quality of data citation behaviour also uncovered that data citation is not carried out adequately with regard to existing requirements of academic journals (Mooney & Newton, 2012) or research data providers (Mahrholz, Reinhold & Rittberger, 2015). Additionally, as argued by Robinson-García et al. (2015), the citedness of research data

¹ This work was supported by the Leibniz Institute for Educational Trajectories (LIfBi).

heavily depends on the quality of data-related information provided by data repositories and varies across disciplines. Furthermore, data citation policies of scientific journals tend to be slightly stricter in the natural sciences than in the social sciences (cf. Blahous et al., 2015). A case study analysing data citation and sharing policies in the environmental sciences also demonstrates that “an overwhelming majority of funding agencies, repositories and journals fail to provide explicit directions for sharing and citing data” (Weber et al., 2011, p. 1). Obviously, making research data accessible and usable is a time-consuming and cost-intensive task. As a consequence, these activities should be appreciated by the scientific community and moreover demand for the inclusion of data citation indicators as a relevant factor in research monitoring. However, in order to make valid statements about data citation quantity and quality, it is necessary to thoroughly analyse the nature of data citation policies within a certain domain. In this paper, data citation policies of eight research data providers in Europe, the United States and Australia within the scope of longitudinal studies in life course research are being evaluated, e.g. with regard to citation principles and sanctions for data users who do not cite adequately. The aim of the study is to outline the different approaches followed by data providers or data repositories in terms of data citation policies which might influence data usage and citation behaviour of researchers in the domain.

Life course research is currently a very dynamic field of research in the social sciences. It provides stakeholders in politics and education with extensive and reliable data about life paths, transitions and decisions in private as well as professional lives. Furthermore, societal changes over extended timeframes of several years or even decades are being monitored. Longitudinal studies in life course research are generally characterized by large sample sizes, different cohorts of participants and various waves of surveys. There is also a strong demand for protecting sensitive personal information, e.g. about performance in school or the parent-child relationship, which are retrieved in these studies at a large scale. As a result, data providers in life course research generally dispose of high data security standards and offer a variety of data access modes, different type of data formats and data granularity. Users generally have to commit to data use agreements and are obliged to cite the research data used according to specific requirements. These data citation policies include aspects of contractual obligations of data citation, concrete requirements of including data citation elements (e.g. a persistent identifier) (cf. Mooney & Newton, 2012) or the position of the data citation within a publication (e.g. in the abstract or the references section) as well as disposal obligations for publications based on the research data provided.

DATA CITATION POLICIES OF DATA PROVIDERS IN LIFE COURSE RESEARCH – A CASE STUDY

For the case study a sample of eight longitudinal studies across the life course in Europe, North-America and Australia was identified by means of six criteria to ensure comparability: 1) thematic focus on educational and personal transitions, 2) ongoing research project, 3) at least a national or international perspective, 4) elaborated data access technologies (e.g. via a data center), 5) data use agreements as a prerequisite for data usage of sensitive data, 6) mention of data citation requirements.² Based on these criteria the following longitudinal studies were selected:

² The criteria were applied to the result set of an extensive web search which retrieved overall 19 longitudinal studies across the life course in Europe, North-America and Australia. The starting point for the web search was a list of longitudinal studies in the social sciences issued by Mallock, Riege & Stahl (2016, p. 146-148).

Table 1. Sample of longitudinal studies across the life course.

Study name	Research topics	Country	Start in year
Étude Longitudinale Française depuis l'Enfance (ELFE)	Impact of family circumstances, living conditions and environment on the physical and psychological development, health and socialization of children.	France	2011
Millennium Cohort Study (MCS)	Influence of early family context on child development and outcomes throughout childhood, adolescence and adulthood.	UK	2000
Negotiating the Life Course	Changing life courses and decision-making processes of men and women as the family and society move from male breadwinner orientation in the direction of higher levels of gender equity.	Australia	1997
National Educational Panel Study (NEPS)	Educational processes from early childhood to late adulthood.	Germany	2009
Panel Analysis of Intimate Relationships and Family Dynamics (pairfam)	Partnership and family dynamics in Germany.	Germany	2008
Socio-Economic Panel (SOEP)	Objective living conditions, values, willingness to take risks, current social changes, and the relationships and interdependencies among these areas.	Germany	1984
Transitions from Education to Employment (TREE)	Post-compulsory educational and labour market pathways of school leavers.	Switzerland	2001
Panel Study of Income Dynamics (PSID)	Employment, income, wealth, expenditures, health, marriage, childbearing, child development, philanthropy, education, and numerous other topics.	US	1968

For each of the research data providers in the domain of longitudinal studies across the life course, the following eight factors were documented by thoroughly eliciting regulatory and user service information on the data providers web sites³: 1) *wording of obligations with regard to data citation*, 2) *requirements for obligatory data citation elements*, 3) *requirements for data citation location within a publication*, 4) *availability of concrete examples for data citation*, 5) *obligation to report data-related publications*, 6) *period of notification for data-related publications*, 7) *disposal obligation⁴ for data-related publications* and 8) *sanctions for*

³ For the analysis, different information sources on the providers' websites were reviewed, e.g. the data use agreements or the specific data citation section. The URLs of the homepages of all data providers in the sample are mentioned in the reference section.

⁴ Publications which are based on a specific dataset are to be submitted to the data provider as a paper-based or digital version according to an agreement of use.

not citing research data in accordance to the requirements. From the point of view that the citedness of research data heavily depends on the quality of data-related information provided by data repositories (cf. Robinson-García et al., 2015), it is legitimate to assume that all of these factors might affect data citation behaviour of researchers in the field.

FINDINGS AND DISCUSSION

All eight data providers issue data use agreements that oblige their users to cite research data. The wording of these obligations (1) in the data use agreements differs significantly, ranging from very concrete citation specifications to rather general requests to cite in accordance to “academic conventions”. Furthermore, all providers name obligatory data citation elements (2): Seventy-five percent of the data providers in the sample demand for including a distinct data version, 50% for including a Digital Object Identifier (DOI)⁵ and 37.5% for naming a specific reference article which outlines the original study design. All eight data providers ask for the inclusion of an acknowledgement phrase indicating either the name of the study or the data center involved. These findings clearly indicate that data providers in life course research generally follow a top-down approach to prevent uncitedness of research data. It is also noticeable that again only 37.5% of providers in the sample provide guidelines for data citation location within a publication (3), e.g. for citing the study as the originator of the data in the title, the abstract or the reference section. This is surprising as it can be assumed that these recommendations are not only useful for guiding data users in the writing process. The recommendations might also foster awareness amongst researchers about the “quality” of a data citation within a document. For example, a data citation in the title or in the abstract can possibly be assessed as more valuable than a data citation in the caption of table or a figure. Interestingly, the data use agreement of the French ELFE study already indicates that users are obliged to cite the study in the title *and* the body of the text if the article is exclusively or primarily based on ELFE data (ELFE, 2014).

Apart from one, all data providers publish concrete examples for data citation on their websites which for example include the names of the authors (of a reference article), the name of the study and the DOI (4). Of course, researchers can already refer to more general data citation guidelines (cf. DataCite, 2014; ESRC, 2016; ZBW, GESIS & RatSWD, 2015⁶). Precise citation examples which relate to the actual study in use might nevertheless be even more important for supporting researchers and help them to prevent citation errors. Seventy-five percent of the providers in the sample insist on the obligation to report data-related publications (5) with only one provider, the Leibniz Institute for Educational Trajectories (LifBi) for the NEPS data, calling for a period of notification for data-related publications of four weeks before publishing (6). And 50% of the data providers even issue a disposal obligation for publications using research data (7). Surprisingly, only one data provider – again LifBi – calls for sanctions if data users do not cite in accordance to the data use agreement (cf. LifBi, 2015) (8)⁷. In summary, it might be assumed that research data providers have already identified a need for action with regard to data citation misbehaviour. However, it still needs to be verified whether the data citation policies described here are

⁵ One data provider has just recently added the obligatory inclusion of a DOI in his citation recommendations – this might be an indicator that the DOI becomes more widely accepted within the domain.

⁶ This publication is not available in English yet.

⁷ Although it is not explicitly stated that non-citations cause a breach of contract, citing the study name and the dataset used for analysis can be interpreted as “essential obligations” of the data use agreement.

appropriate measures for achieving high citation rates and citation quality of research data issued by providers of longitudinal data in life course research.

As stated in the introduction, the main goal of the study was to outline the variety of data citation policies within life course research and to discuss possible implications for data use and citation behaviour in the field. It could be demonstrated that data providers follow differing approaches in terms of data citation requirements. This involves data versions, identifiers and reference articles describing the original study design. In addition, data providers differ substantially with regard to recommendations for data citation location as well as disposal obligations for data-related publications. This might lead to a high diversity in data citation behaviour of data users in the field and potentially to non-comparable results in data citation analysis. It is therefore reasonable to argue that data providers should pursue the harmonisation of data citation specifications – in close cooperation with journals and research institutions involved in life course research. Furthermore, policy makers should strongly encourage the development of domain-specific data citation indicator sets for the valid representation of scientific output, allowing for an improved comparability and traceability of research.

LIMITATIONS OF THE STUDY AND OUTLOOK

We are aware that our research has some limitations. First, the study consists of a small sample which is not representative for data usage and citation within the social sciences in general. Second, there might be other longitudinal studies in life course research that meet the selected criteria presented above. Third, there is a predominance of European longitudinal studies in the sample. Finally, the study does not investigate the influence of data citation policies on the actual data citation behaviour of researchers in the field. A consecutive study, analysing data citation quantity and quality in a large sample of data-related publications in the social sciences might substantially enhance our understanding of data citation behaviour. Despite these limitations we believe our work has highlighted the importance of critically examining data citation policies beforehand as one milestone of coherent and comparable data citation analysis.

REFERENCES

Blahous, B., Gorraiz, J., Gumpenberger, C., Lehner, O., Stein, B. & Ulrych, U. (2015). Forschungsdatenpolicies in wissenschaftlichen Zeitschriften – Eine empirische Untersuchung. In *ZfBB*, 62, 12-24.

DataCite International Data Citation Metadata Working Group (2014). DataCite Metadata Schema for the Publication and Citation of Research Data. Version 3.1 October 2014. Retrieved March 16, 2016 from: http://schema.datacite.org/meta/kernel-3/doc/DataCite-MetadataKernel_v3.1.pdf.

Étude Longitudinale Française depuis l'Enfance (ELFE) (2014): Charte d'accès aux données Elfe. Retrieved March 16, 2016 from: http://www.elfe-france.fr/images/documents/charte_acces_donnees_plateforme.pdf.

Economic and Social Research Council (ESRC) (2016). Data Citation. What you need to know. Retrieved March 16, 2016 from:

https://www.ukdataservice.ac.uk/media/104397/data_citation_online.pdf.

Leibniz-Institute of Educational Trajectories (LifBi) (2015). NEPS. Datennutzungsvertrag.

Retrieved March 16, 2016 from: <https://www.neps-data.de/Portals/0/NEPS/>

[Datenzentrum/Datenzugangswege/Vertraege/NEPS_Datennutzungsvertrag_DE.pdf](https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Datenzugangswege/Vertraege/NEPS_Datennutzungsvertrag_DE.pdf).

Mahrholz, N., Reinhold, A. & Rittberger, M. (2015). Data Citation Quantity and Quality in Research Output of a Large Educational Panel Study. In *Proceedings of the 15th International Conference of Knowledge Technologies and Data-Driven Business (I-Know '15)*. Graz: ACM.

Mallock, W., Riege, U., Stahl, M. (2016). *Informationsressourcen für die Sozialwissenschaften. Datenbanken, Längsschnittuntersuchungen, Portale, Institutionen*. Wiesbaden: Springer.

Millennium Cohort Study (MCS). Retrieved March 16, 2016 from:

<http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=851&sitesectiontitle=Welcome+to+the+Millennium+Cohort+Study>.

Mooney, H. and Newton, M. P. (2012). The Anatomy of a Data Citation: Discovery, Reuse, and Credit. In *Journal of Librarianship and Scholarly Communication*, 1, 1, eP1035.

Mooney, H. (2011). Citing data sources in the social sciences: do authors do it? In *Learned Publishing*, 24, 2, 99-108.

Negotiating the Life Course. Retrieved March 16, 2016 from:

<http://lifecourse.anu.edu.au/>.

Panel Analysis of Intimate Relationships and Family Dynamics (pairfam). Retrieved March 16, 2016 from: <http://www.pairfam.de/en/>.

Panel Study of Income Dynamics (PSID). Retrieved March 16, 2016 from:

<https://psidonline.isr.umich.edu/>.

Peters, I., Kraker, P., Lex, E., Gumpenberger, C. & Gorraiz, J. (2015). Research Data Explored: Citations versus Altmetrics. In *Proceedings of the 15th International Conference on Scientometrics and Informetrics (ISSI '15)* (p. 172-183). Istanbul: Bogaziçi University Printhouse.

Piwowar, H.A., Carlson, J.D. & Vision, T.J. (2011). Beginning to track 1000 datasets from public repositories into the published literature. In *Proceedings of the American Society for Information Science and Technology*, 48, 1, 1-4.

Piwowar, H.A., Day, R.S. & Fridsma, D.B. (2007). Sharing Detailed Research Data Is Associated with Increased Citation Rate, *PLoS ONE* 2, 3, e308. Retrieved March 16, 2016 from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1817752/>.

Robinson-García, N., Jiménez-Contreras, E. & Torres-Salinas, D. (2015). Analyzing data citation practices using the data citation index. In *Journal of the Association for Information Science and Technology*, DOI: <http://dx.doi.org/10.1002/asi.23529> (Preprint).

Socio-Economic Panel (SOEP). Retrieved March 16, 2016 from: <http://www.diw.de/en/soep>.

Transitions from Education to Employment (TREE). Retrieved March 16, 2016 from http://www.tree.unibe.ch/index_eng.html.

Weber, N.M., Piwowar, H.A. & Vision, T.J. (2010). Evaluating Data Citation and Sharing Policies in the Environmental Sciences. In *Proceedings of the American Society for Information Science and Technology*, 47, 1, 1-2.

ZBW, GESIS & RatSWD (2015). Auffinden, Zitieren, Dokumentieren: Forschungsdaten in den Sozial- und Wirtschaftswissenschaften. Retrieved 16 March, 2016 from: http://auffinden-zitieren-dokumentieren.de/wpcontent/uploads/2015/03/Forschungsdaten_DINA4_ONLINE_VER_02_06.pdf.



Stepping up Information Infrastructures and Statistical Reporting – Monitoring the German Excellence Initiative

Anke Reinhardt

*anke.reinhardt@dfg.de

Information Management Unit, Deutsche Forschungsgemeinschaft (DFG, German Research Foundation),
Kennedyallee 40, 53175 Bonn (Germany)

ABSTRACT

The Excellence Initiative has not only been the most prominent funding scheme in German research policy in recent years, but has also had important side effects on research management. This paper argues that the Excellence Initiative was indeed a “boost” for improving the data infrastructure and statistical reporting of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation). The learning effects are now transferred to the line business and serve as a good starting point for the reporting on a potential third phase of the Excellence Initiative.

BACKGROUND

The Excellence Initiative is a funding scheme launched in 2005 with the aim of promoting top-level research in Germany. It has three funding lines: “Institutional Strategies for Top-Level University Research” are supposed to increase the international competitive ability of the entire university; “Graduate Schools” (GSC) should provide highest-level research training; “Clusters of Excellence” (EXC) pool excellent researchers in interdisciplinary centres.

The DFG is the largest funding agency for basic research in Europe, supporting almost 30 000 research projects in all scientific disciplines (Deutsche Forschungsgemeinschaft 2015a). Bund and Länder jointly commissioned the DFG – together with the Wissenschaftsrat (WR, German Council of Science and the Humanities) – to implement the funding scheme on their behalf, and with an additional budget.

So far, the Excellence Initiative runs until 2017 for two funding periods (2006-2012 and 2012-2017). After DFG and WR provided a report about the implementation of the Excellence Initiative in 2008 (Sondermann et al. 2008), Bund and Länder granted a second round of funding. However, they also agreed that a decision about a renewal of the Excellence Initiative after 2017 should be based on an external evaluation (“Imboden Commission”). To support this evaluation, in 2015 DFG and German Council of Science and Humanities (WR) were again expected to report on the course of the Excellence Initiative, this time in a “data-based way”.

the challenge of the “data-based report”

The purpose of the data-based report was to provide information on effects and – if possible – output of the funded projects. Of particular interest were structural effects, e.g. the working conditions of PhD candidates, international appeal and recruitment successes of universities, interdisciplinarity and cooperation ties with non-university research institutes or industry.

The usual source for statistical services of the DFG is its internal database, which stores process-produced data on 620,000 applications (e.g. requested resources, discipline, time span, collaborators), on applicants and reviewers (gender, age, nationality etc., in total 230,000 scientists) and on 50,000 research institutes. Besides the actual grant management, the DFG uses this data to steer the programmes and to perform detailed analyses on all aspects of its funding, e.g. success-factors (Dinkel & Wagner 2015), participation of women (DFG 2015b), internationalization (Fuß 2011) or interdisciplinarity (Güdler 2013). For in-depth analysis, it commissions evaluation studies, which also include additional data sources, e.g. interviews, bibliometrics, surveys etc. (Reinhardt 2013).

The DFG, being responsible for the data-based reporting on EXC and GSC, had to focus much more on throughput and output information than it usually does. It decided to include other data sources in addition to its internal database. Additionally, it needed a different analytical perspective on the effects of science funding, with the particular difficulty that almost all universities in Germany participated in the Excellence Initiative.

NOVELTIES IN THE DATA INFRASTRUCTURE AND LINES OF ANALYSIS

To develop the concept of the data-based report, the DFG established a new organisational setting: It consulted with a scientific advisory board of five eminent scientists in the field of research evaluation and science of science studies on the question which data to collect and how to analyse it.

One measure was that the DFG started to collect data on the “life” in projects. EXC and GSC have a large number of scientific members (usually between 100 and 600) not involved in the application for funding, who are therefore “unknown” to the DFG data-base. The DFG collected information on the doctoral candidates, the postdocs, research group leaders, guest researchers and other scientific staff, together more than 20,000 persons, as well as on professorial positions created.

Additionally the DFG hired a contractor to analyse renewal applications of GSC and EXC and of the protocols of the peer review panels to identify organisational measures and structural effects mentioned there. An online survey of 990 involved researchers as well as interviews and focus groups, mainly with Principal Investigators and presidents of universities, asked about experiences and opinions of the Excellence Initiative. A survey of reviewers complemented the picture (Möller 2012).

The report to Bund and Länder summarised information from all these sources. Additionally to this report, the DFG published analyses that dig deeper into specific effects.

One of the most prominent reporting products of the DFG is the Funding Atlas (DFG 2015c). It provides information on public research funding in Germany, particularly DFG funding, at German universities. The 2015 edition specifically looked at concentration effects of the Excellence Initiative on funding and on disciplinary profiles of universities using the Gini coefficient. It turns out that the Excellence Initiative did not increase the concentration but instead leveraged more grant-seeking activity at all German universities. It strengthened subjects that were strong before. Network analysis of disciplines, based on the classification of proposals, and on regions, based on researcher’s location, deepened the understanding of

the German research landscape. Additionally a bibliometric analysis of two subjects indicates that the funded universities were slightly more productive than others ¹.

DISCUSSION

The political importance attached to the Excellence Initiative proved to be a catalyst for change. The DFG used some of the novelties introduced in the Excellence Initiative to step up its data infrastructure and reporting more generally. For example, it transformed the survey instrument introduced in the Excellence Initiative to also survey the CRC and RTG, which in the future allows to compare these funding lines. Equally, it uses the document analysis methodology used in the Excellence Initiative in a project analysing the effects of its “Research Oriented Standards on Gender Equality”.

After the Excellence Initiative is before the Excellence Initiative: Currently politics debates about the shape of a future round. While it has already agreed on an extension, the exact format will only be decided in June 2016. However, it seems that in the future even more focus will lay upon the effects and side-effects: Will the large number of PhDs educated in the GSC be able to find qualified jobs? Are the recruited researchers there to stay? Has the governance of universities changed for good? Answering these questions requires other kinds of data and other approaches than the ones used previously.

The DFG will therefore place even more emphasis on output data. Final reports are a good source of information on publications, scientific content and staff. The DFG plans to start collecting final reports electronically which allows to analyse the information more easily. A specific focus will be on the text analysis of proposals, on tracking research topics, and on career outcomes by researching the placement of staff members.

In the meantime, new data sources are available. For example, funding acknowledgements allow links between funding and specific publications as well as their citation rates. The DFG needs to enforce its policy on this. A new statistics on doctoral researchers by the German Federal Statistical Office can supplement information that the DFG collects in its survey.

The more universities are involved in the Excellence Initiative, the harder it will become to single out the effects of the Excellence Initiative. The DFG is therefore eager to cooperate again with researchers to ask the rights questions about this funding instrument and to test novel methodologies of analysis. This will allow gaining a deeper understanding not only on the Excellence Initiative but on research funding more general.

¹ The DFG was not alone in analysing the effects of the Excellence Initiative. For example, the Berlin-Brandenburg Academy of Sciences and the Humanities published a bibliometric report on the relative success of universities funded by the Excellence Initiative versus others in terms of publications and citations (Hornbostel & Möller 2015). Engels et al. analysed its effects on Gender equality (Engels et al. 2015).

REFERENCES

- Deutsche Forschungsgemeinschaft (2015a): Jahresbericht 2014, Bonn: DFG.
www.dfg.de/dfg_profil/jahresbericht
- Deutsche Forschungsgemeinschaft (2015b): Chancengleichheits-Monitoring 2014. Bonn: DFG.
www.dfg.de/en/research_funding/principles_dfg_funding/equal_opportunities/monitoring_equal_opportunity
- Deutsche Forschungsgemeinschaft (2015c): Förderatlas 2015. Kennzahlen zur öffentlich finanzierten Forschung in Deutschland. Bonn: DFG.
- Dinkel, W. & J. Wagner (2015): Who applies for DFG grants? Statistical Trends in Proposal Submission and Applicant Demographics. *DFG Infobrief*, 1. Bonn: DFG.
www.dfg.de/en/dfg_profile/facts_figures/evaluation_studies_monitoring/info_briefe
- Engels, A. et al. (2015): [Bestenauswahl und Ungleichheit](#). Eine soziologische Studie zu Wissenschaftlerinnen und Wissenschaftlern in der Exzellenzinitiative. Frankfurt a.M.: Campus.
- Fuß, K. (2011): Doctoral Training Frameworks. A Comparison of Doctorates in the DFG's Research Training Groups and Collaborative Research Centres Programmes. *DFG Infobrief*, 4. Bonn: DFG.
www.dfg.de/en/dfg_profile/facts_figures/evaluation_studies_monitoring/info_briefe
- Güdler, J. (2013): Interdisciplinary review processes: structural impact and funding success. Bonn: DFG.
www.dfg.de/dfg_profil/zahlen_fakten/evaluation_studien_monitoring/studien/bericht_interdisziplinaritaet/
- Hornbostel, S., & Möller, T. (2015): Die Exzellenzinitiative und das deutsche Wissenschaftssystem: Eine bibliometrische Wirkungsanalyse. *Wissenschaftspolitik im Dialog 12*. Berlin: Berlin-Brandenburgische Akademie der Wissenschaften.
- Möller, T. et al. (2012): Exzellenz begutachtet. Befragung der Gutachter in der Exzellenzinitiative. *iFQ-Working Paper*, 11. Berlin: iFQ.
- Reinhardt, A. (2013): Das Evaluationsvorgehen der Deutschen Forschungsgemeinschaft. *Zeitschrift für Evaluation*, 2, pp. 305-314.
- Sondermann, M. et al. (2008): Die Exzellenzinitiative. Beobachtungen aus der Implementierungsphase. *iFQ-Working Paper*, 5, Bonn: iFQ.

CHAPTER 2

Smart Use of Indicators for Innovation Policy



Innovation indicators: Towards a User's guide

Michiko Iizuka* and Hugo Hollanders**

*iizuka@merit.unu.edu

UNU-MERIT, Boschstraat 24, Maastricht, 6211AX (The Netherlands)

** h.hollanders@maastrichtuniversity.nl

UNU-MERIT, Boschstraat 24, Maastricht, 6211AX (The Netherlands)

ABSTRACT

The increased use of innovation indicators is observed in policy sphere. Several reasons are attributed for this increased use: first, access to data on innovation increased as the result of over half a century of efforts made by international organizations and researchers in this field; second, an increase in types and number of innovation indicators; third, its increased use in business contexts with the understanding that innovation is an integral part of business performance, which further spurred the use and development of innovation indicators (Soete and Freeman, 2009); fourth, the expansion in geographical coverage of countries, particularly in emerging countries of Latin America, Asia and Africa (Gault, 2010) making innovation indicators more policy (politics) relevant in a wider variety of countries, e.g. by allowing to benchmark and compare countries' innovation performance; and fifth, the innovation indicators (in particular, the composite indicator but any indicator if used in ranking countries) became a communicative tool in public debate, in the backdrop of recent emphasis on 'evidence based policy' and 'participatory decision making' in the policy domain (OECD, 2012).

Considering innovation indicators are intended to improve the performance of innovation policy, their increasing use is generally good news. Nevertheless, due to the increase in diversity in type and context in which indicators are being applied there is an increased occurrence of inappropriate use and misinterpretation of innovation indicators in the policy sphere. This paper first describes the specific cases then tries to generalize the problem aiming to build a general guideline or check list on the appropriate use of innovation indicators.

1. MOTIVATION/INTRODUCTION

The increased use of innovation indicators is observed in policy sphere. Several reasons are attributed for this increased use: first, access to data on innovation increased as the result of over half a century of efforts made by international organizations and researchers in this field; second, an increase in types and number of innovation indicators; third, its increased use in business contexts with the understanding that innovation is an integral part of business performance, which further spurred the use and development of innovation indicators (Soete and Freeman, 2009); fourth, the expansion in geographical coverage of countries, particularly in emerging countries of Latin America, Asia and Africa (Gault, 2010) making innovation indicators more policy (politics) relevant in a wider variety of countries, e.g. by allowing to benchmark and compare countries' innovation performance; and fifth, the innovation indicators (in particular, the composite indicator but any indicator if used in ranking countries) became a communicative tool in public debate, in the backdrop of recent emphasis on 'evidence based policy' and 'participatory decision making' in the policy domain (OECD, 2012).

Considering innovation indicators are intended to improve the performance of innovation policy, their increasing use is generally good news. Nevertheless, due to the increase in diversity in type and context in which indicators are being applied there is an increased occurrence of inappropriate use and misinterpretation of innovation indicators in the policy sphere. This paper first describes the specific cases then tries to generalize the problem aiming to build a general guideline or check list on the appropriate use of innovation indicators.

2. PROBLEMATIC USE OF INNOVATION INDICATORS IN THE POLICY DOMAIN

Some of the problematic use of innovation indicators is categorized into five subheadings. The first problematic use concerns the misinterpretation of innovation indicators due to a basic misconceived notion associated with innovation, such as R&D is a precondition for innovation, more or higher performance on a particular innovation indicator is always better and more innovation lead to positive outcomes. The second problem concerns the inappropriate use of innovation indicators. The two problems listed under this heading are similar in adapting a simplistic view of a complex reality: one problem concerns the construction or design of an indicator and the other problem the context in which it is applied. In another words, an indicator is often used without a clear understanding of its construction (design) or even if the user understands the design of the indicator, the user does not know whether the existing design can be used in a different context (be it industry, sector or country). The third problematic use concerns the misuse of an innovation indicator in the policy domain. This problematic use is very much related to the two before-mentioned subheadings, misinterpretation and inappropriate use, but is more strongly linked to the use for policy purposes. The fourth problematic use concerns mainly the 'unknown' part of an innovation indicator due to the current 'omission' of some of the factors selected to be used as 'indicator'. The selection of relevant indicators to get a grasp of a complex and multi-dimensional reality such as innovation is a difficult task. Furthermore, as innovation is a dynamic process, some factors selected as indicators can become obsolete in the new context while other factors gain importance but it is simply 'not known' at the moment (an example would be the ownership of mobile phone or access to internet). The last problematic use concerns the issue of practices regarding the use of indicators in particular addressing the needs or use of indicators. Indicators are made to be used for improving innovation policy but sometimes there are mismatches in how indicators are designed and delivered and how indicators were demanded to be used and delivered.

In the following section, the problematic issues of innovation indicators in policy use are discussed by type of innovation indicators categorized as follows: (1) Traditional indicators used as innovation indicators (such as R&D expenditures, Patent applications, Human Resources and Scientific publications); (2) Innovation indicators collected through surveys; (3) Composite indicators.

BOX1 The ‘problematic’ use of innovation indicators in policy domain identified

1. Misinterpretation of innovation indicator

Assuming linear progression of innovation that R&D precedes innovation

Assuming that more/higher performance on an indicator is always better (Foray and Hollanders, 2015);

Assuming that more innovation automatically leads to development (Soete, 2013).

2. Inappropriate use of innovation indicator

Compare indicators, which are not comparable (due to different collection methods, assumptions, measurements, industrial structure);

Applying same indicator criteria in different sector and country setting without careful consideration of characteristics or new contexts in which the indicators are being applied.

3. Misuse of innovation indicator for policy purposes

Blindly applying the R&D/GDP target (in developing countries, it is 1% and in developed countries, it is 3%) as a policy goal without understanding a country’s industrial structure and HR composition;

Applying the indicator to policy formulation without understanding the underlying conceptual design and data collection procedure;

Ignoring the country/sector/industrial structural context when interpreting the innovation indicators for policy purposes;

Relying only on composite indicator ranking to monitor, evaluate and formulate innovation policy (and to make political statements that would mislead the public).

4. Overlooked issues of innovation indicators in policy domain

Omitting important sources of innovation which are actually vital for the economy and therefore for policy formulation (in developed context, e.g. vocational education as an alternative to tertiary education (Foray and Hollanders, 2015) Globalization of business activities in developed context (Edquist and Zabala-Iturriagagoitia, 2015), Non R&D oriented innovation (i.e user led innovation/household innovation, public sector innovation?), in developing countries, informal R&D (user led innovation/household innovation), those who are trying to innovate from those who do not do anything, External sources of knowledge (Foreign Direct Investment (FDI), exports, GVC), informal sector, other productive sectors important in developing countries).

Ignoring the dynamic nature of industrial structure and relevance of selected indicators.

5. Mismatch of needs between user and producers of innovation indicator

a. Ignoring the results of innovation survey for policy elaboration (because the data come late (questions about its usefulness))

b. Ignoring the importance of comparability for the indicators (changing questions not to add).

Source: authors

3. THE PROPORTION OF RESEARCH AND DEVELOPMENTAL EXPERIMENTATION (R&D) IN GDP

R&D data are the most available data for the longest period of time and covers many countries. The concept and specification of R&D statistics are defined by the Frascati manual. Basically, R&D covers basic research, applied research and experimental development. This definition appears relatively straightforward but when data on R&D are to be collected it is not easy to distinguish research activities from non-research activities because the distinction between these two is determined by the ‘intention’ of actions taken. For example, ‘action of taking temperature measurement’ can be categorized as research and experimentation if the ‘intention’ was for research and experimentation while if the purpose was ‘routine’ activities, it is not counted as R&D activities.

Share of R&D in GDP is often used as a policy guideline to improve the innovation policy. For instance, targeting a certain percentage of R&D spending in GDP is often used as the policy goal. In fact, for the European Union, the Lisbon agenda sets 3% as the target; while many African and Latin American countries have 1% as their goal. While this can be useful as a general guideline, blindly applying the target to different country contexts, assuming that a higher percentage of GDP spent on R&D would lead to development misleads innovation policy. The reasons for such are as follows:

Appropriate to industrial structure

First, R&D intensities differ across industrial activities. Countries with different industrial structures should have different levels of the percentage of R&D that is appropriate for a given industrial structure. In another words, policy makers should pay attention to the efficiency and match of R&D expenditure to the needs of their country’s industrial sector not just to increasing the share of R&D in GDP.

For instance, the OECD classifies the types of industries by the R&D intensities (e.g. in terms of R&D as a share of value added, R&D as a share of production, R&D plus technology embodied in intermediate and investment goods as a share of production). Currently the OECD uses a four-tier model to classify industries with R&D intensities are follows:

Box 2 R&D intensity

	Direct + indirect R&D as a share of production	R&D as a share of production	R&D as a share of value added
High tech industries	Above 7.5%	Above 7.5%	Above 15%
Medium high tech industries	Between 2.5% and 7.5%	Between 1.5% and 7.5%	Between 4% and 15%
Medium-low tech industries	Between 1% and 2.5%	Between 0.5% and 1.5%	Between 1.5% and 4%
Low tech industries	Below 1%	Below 0.5%	Below 1.5%

Source: Hatzichronoglou, 1997

Note: High tech industries include Aircraft and spacecraft, Pharmaceuticals, Office, accounting and computing machinery, Radio, TV and communications equipment, Medical, precision and optical instruments. Medium-high technology industries include Electrical machinery and apparatus, Motor vehicles, trailers and semi-trailers, Chemicals excluding

pharmaceuticals, Railroad equipment and transport equipment, Machinery and equipment. Medium-low technology industries include Building and repairing of ships and boats, Rubber and plastics products, Coke, refined petroleum products and nuclear fuel, Other non-metallic mineral products, Basic metals and fabricated metal products. Low tech industries include Manufacturing, Recycling, Wood pulp paper, paper products printing and publishing, Food products, beverages and tobacco, Textiles, textile products, leather and footwear.

The categorization of technological level had slightly changed recently (2011) but overall principle of associating the type of activities (ISIC code) to the level R&D intensity continues (see technical note, OECD, 2011 and Hatzichronoglou, 1997).

Countries with higher shares of high tech industries are more likely to have higher share of R&D in GDP while countries with higher shares of medium-low tech industries (like Southern European countries) would have lower shares of R&D in GDP. In this way, appropriate level of the ‘optimal’ share of R&D in GDP can differ due to the industrial structure of a country. Hence a 3% or 1% guideline should be taken only as the guideline and not to be applied blindly in policy.

Bias towards manufacturing/high tech sector

Second, by design, R&D measurements are highly biased towards the manufacturing sector. This would create a problem when different sectors such as service, agriculture and natural resource based activities are to be assessed applying the same methods. This point is already being identified by the OECD. The technical notes of OECD directorate for STI states that “Direct R&D intensities are not much help for service activities. Instead other indicators such as skill intensity (e.g. education levels in industry x occupation matrices) and indirect R&D measures such as technology embodied in investment or investment in ICT goods by industry must be explored.” (OECD, 2011). The same document also admits the limitation in disaggregating low tech industries due to the limited detailed R&D expenditure data across countries. On low tech industries, several studies also question the underlying assumption associated with low tech and low knowledge/technology intensity (Hirsch-Kreinsen and Schwinge, 2014, von Tunzelmann and Acha, 2005, Mendoca and von Tunzelmann, 2004). Hence, applying the preconceived notion from a particular context cannot be applicable to measure the conditions in different countries. In another words, the proportion of R&D in GDP cannot be used as the sign for the innovativeness of a country.

Different policy implication due to the origin of R&D funding

Third, differences in the origin of R&D (public versus private) are another distinction that needs to be considered. In general, developed countries have larger proportions of R&D performed and financed by the private sector while in developing countries the major contribution to R&D is made by the public sector. This difference will have different policy implications. In countries where the private sector is more active in R&D, policies targeting the private sector (policies such as tax incentives, subsidies etc.) can boost the share of R&D in GDP by raising business R&D expenditures. If the share of R&D is larger in the public sector, then increasing the R&D would need to be preceded with policies to enhance human resources to carry out R&D and investment in the public research infrastructure (laboratories, university and research institutions, administrative capacities to carry out R&D).

More firms conduct innovation than R&D

One of the misinterpretations that is easily identified in developing countries is the assumption of linear progression that innovation always comes after R&D. As evidenced by European survey data, about half of European firms that innovate do not conduct R&D (Huang, Arundel, Hollanders, 2007). The share of firms that innovate without doing R&D is likely even higher in developing countries, where much of the early challenge is to deal with existing ‘bottlenecks’ (Sutz, 2012) or ‘weak innovation systems’ (UNU-INTECH, 2005). Policies in developing countries should therefore pay sufficient attention to innovation in terms of organizations, non-technological innovation and the import of embodied technologies not involving own R&D activities.

4. INNOVATION SURVEYS (CIS, FOLLOWING OSLO MANUAL)

Innovation surveys are conducted to collect information on innovation. Innovation surveys, in Europe represented by the Community Innovation Survey (CIS), follow the Oslo Manual (3rd revision) guidelines how to measure innovation. Innovation surveys ask the performers of innovation (i.e. firms) whether they conduct certain activities that lead to innovation. The definition of innovation, collection methods, survey questions, and data compilations have evolved over the years to improve the quality of statistics and it is closely linked with the evolutionary change in the Oslo Manual (from original to revision 4)¹. The survey developed for European countries, the CIS, and the Oslo Manual are applied in most of the emerging countries by adapting the questionnaire to the local context while keeping comparability. The degree of modification of the CIS questionnaire essentially depends on the choices of these countries on what they want to find out regarding innovation and innovation policy.

Innovation survey data basically complement existing data on patents, bibliometric indicators and R&D surveys. Hence, the survey basically provides the following information (Mohnen and Mairesse, 2010: 6):

- Indicators of innovation output (such as the introduction of new products and processes, organizational changes and marketing innovations, the percentages of sales due to products new to the firm or new to the market, and the share of products at various stages of the product life-cycle);
- A wider range of innovation expenditures or activities than R&D expenditures (such as the acquisition of patents and licenses, product design, personnel training, trial production, and market analysis);
- Information about the way innovation precedes, such as sources of knowledge, the reasons to innovate and perceived obstacles to innovation.

The Oslo Manual follows the subject approach of survey which is collecting information from the firm level instead of object approach, which collects information on innovation (SPRU study), the number of innovation ‘output’. The subject approach collects comprehensive data at the decision making level of the firm allowing to conduct much richer analysis that can be linked to the sectoral statistics and national accounts while the drawback of the subject approach is that it does not distinguish between successful and unsuccessful innovations.

The data obtained from innovation surveys are qualitative, subjective and censored (Mohnen and Mairesse, 2010). The number of variables are censored and selected as samples (unless

¹ Detail history of evolutionary development please see following: for Oslo manual (Gault, 2013) and for Community Innovation survey (Arundel and Smith, 2013).

otherwise it is census) and hence subject to some biases (for example, sector). The information obtained is subjective and the quality of variables may contain errors.

Developing countries' problem of adapting the survey:

Innovation indicators started to being adapted in many developing countries since the 1990s. In fact, implementation of innovation surveys in Latin America is not so different from that of Europe (RYCT) and similar interests were also expressed in Africa as can be seen from a NEPAD study (UNU-INTECH). While the recognition on the importance of innovation was present from the early days, earlier experiences of applying Oslo manual based innovation surveys suffered difficulties in not quite capturing the particularities of developing countries. The Bogota manual, as the result, was produced by Colciencia in response to meet the different ideosyncrasy of the Latin American innovation context which were later incorporated in the annex of the third revision of the Oslo manual. Many developing countries are currently trying to start conducting innovation surveys. Most of these countries follow the Oslo manual by adapting the CIS survey to understand the innovation process in the country (Gault, 2013, Crespi and Periano, 2007). Many developing countries question the usefulness of conducting an innovation survey. The reasons are as follows.

High cost and barrier

In developing countries, collecting data is much more difficult due to not having fully equipped and capable statistical offices, who may need to prioritize different demands coming from the government (be it demographic data, household survey data etc). In other words, the opportunity cost of conducting an innovation survey is high, especially compared to developed countries. Some of these countries may need to start from building business registries to have acceptable level of selectivity.

Fitting to its economic and industrial structure?

Furthermore, as innovation surveys were originally designed for the developed countries, survey results may not reflect the actual economic/industrial reality in developing countries. For example, many developing countries have a large informal sector (de Beer et al, 2013, Iizuka et al., 2015, Konte and Ndong, 2012). This means that even with well-developed business registries, the survey can only illustrate a relatively small part of economic activities. Moreover, even if the survey is conducted following the Oslo manual, with guidelines based on the experiences of developed countries, copy-pasting the survey questions would not lead to the output that may serve the needs of policy makers in improving innovation policy. For instance, the industrial structure of many African countries demonstrate the important role played by agriculture in its contribution to economic activities as well as in creating employment (see table in Iizuka et al, 2015). The CIS and Oslo manual currently cover the manufacturing, service and mining and quarry sectors, however, they do not cover agriculture. Hence, survey methodologies that can capture the innovation process in agriculture is needed. In fact, nascent attempts are made in Agriculture by ANNI in Uruguay where they have surveyed the agricultural sector (Aboal et al, 2015). While these attempts were already being made, it could take a rather long time to standardize survey questions to be shared among countries with a large agricultural sector.

Finer adjustment to how developing countries innovate

In the similar vein, some of the questions typically used for innovation surveys may require an adaptation to the reality of developing countries. For instance, the minimum size of the firm to be surveyed would be much smaller in developing countries. The definition and type

of ‘innovative activities’ in developing countries should include (Sutz, 2012) for instance, acquisitions of embodied technology (equipment), minor or incremental changes made in production process, organizational changes, and intentions to conduct innovation.

Sutz (2012) also states that as many developing countries are still yet to develop Innovation capabilities, the investments made in building the system should also be considered as part of expenditure. This would involve the investments in human resources, linkages, quality assurance systems and use of ICTs (Intarakumnerd, 2007).

How to make sense of the survey results to relevant actors?

Resources are often limited in developing countries but some countries manage to conduct an innovation survey. While this is good news, countries are often confronted with other problems: applying the information for policy purposes. For example, the existing survey conducted by the ANII (Uruguay), demonstrated a very low share of policy makers had actually used survey results for innovation policies (Baptista et al. 2009). A comparative study among Chile, Colombia and Uruguay showed similar tendencies. Possible reasons for low policy use of innovation survey data are as follows. First, innovation survey data only become available after some time (results becomes available one year (if not more) later than the reference year of the survey) so it is likely that these data are not perceived as ‘up to date’ enough to be readily used for policy making. Second, these data may not be elaborated in the way policy makers can comprehend and use them correctly. Third, restricted availability and accessibility of data (in particular micro or firm level data) may cause an insufficient analysis of the data (this may lead to the question of making data publically accessible taking into account confidentiality issues).

Furthermore, considering the globalization of activities through extending value chains, many developing countries are technologically catching up through entering markets by producing goods at lower prices. In many developing countries less patentable ‘process’ incremental and organizational innovations would be more prevalent than radical innovations through active investments in R&D.

Thirdly, the structural composition of developing countries should be considered carefully. The trends of developing countries are diverse. For instance, many African and Latin American countries have industrial structures with less diversity and reliance on natural resources while some had experienced strong growth in services. In addition to above differences, the size of the informal economy is also substantial in these countries (de Beer et al, 2013, among others).

For instance, in developing countries where most of the countries do not conduct R&D to innovate (Gault, 2010), “learning” and “problem solving” are important parts of the innovation process. In developing countries, due to the under provision of various basic infrastructures (physical, legal, institutional), much of firms’ innovative efforts are being made in overcoming existing ‘problems’ which are not directly considered as ‘innovation’ in a developed country context (Sutz, 2012). Hence, more firms conduct innovations that do not have R&D nor involve new technology in developing countries. The above example demonstrates the presence of the gap in what constitutes ‘problem solving’ and ‘learning’ in different context even though the same word is used.

5. CASE OF COMPOSITE INDICATOR

Due to an increasing availability and accessibility of diverse sets of data, composite indicators are more and more easily constructed and used in the policy domain. Composite indicators summarize individual indicators by compiling these into a single index. Several composite indicators to measure ‘innovation’ capacity at country level emerged recently such as the Global Innovation Index (WIPO), Global Competitiveness Report (World Economic Forum) and the European Innovation Scoreboard (European Commission) just to name a few. The use of composite indicators became prevalent in the 2000s. Due to the ability of a composite indicator to summarize multidimensional characteristics of complex ideas such as innovation and its facility to communicate and compare results, composite indicators are a powerful policy tool by creating a policy narrative (Saltelli, 2007) while caution of these users are well expressed by numerous experts (OECD JRC handbook, 2008, Freudenberg, 2003, Nardo and Saisana, 200x, Foray and Hollanders, 2015, Edquist and Zabala-Iturriagoitia 2015, Shibany and Streicher, 2008; Adam, 2014 amongst others).

While the intended use of composite indicators is to grasp overview and monitor progress for policy purposes (Grupp and Mogege, 2004), the ranking table of indices is easily being politicized and a powerful tool (for policy makers to dialogue with the public/budget officers) to mobilize the policy agenda by creating a narrative (Saltelli, 2007). While policy makers can use composite indicators to comply with ‘evidence based’ and ‘participatory’ policy making requirements that are increasingly being presented, many users of these indicators may not have a clear understanding how these indicators are constructed and the limitations in what can be interpreted. This potentially creates the information and knowledge asymmetry between different types of users (e.g. policy makers, academics, journalist, lay citizen) making both intentional and unintentional misuses of composite indicators possible. Composite indicators, by definition, give a relative performance benchmark between countries. A common mistake is that a decline in rank performance is interpreted as a real performance decline whereas in most cases a lower rank is not the result of a declining performance but of other countries’ performance improving at an even faster rate.

This is especially true in identifying policy prescriptions using composite indicators. For instance, the composite indicator should be analyzed with alternative data sets to understand in detail about the country in disaggregate form. E.g., two countries can have identical scores for their composite indicator hiding significant differences on some of the underlying pillars with one country clearly performing better on inputs in the innovation process like human resources and the other country on outputs of the innovation process like exporting knowledge-intensive products or selling technological knowledge (technological balance of payments). For such diagnostic purposes, innovation survey data and other available information on R&D, human resources and economic indicators become useful. In fact, several studies (OECD/JRC, 2008, Nardo and Saisana, 2008, Adams, 2014) clearly stated, composite indicators are good in evaluating a country’s innovation performance in relative terms compared to other countries on selected indicators for well-defined purposes; however, these indicators are not well fitted to conduct policy analysis for evaluating and monitoring the implemented innovation policies in detail.

6. CONCLUSION: TOWARDS DISCUSSION

In interpreting innovation indicators and composite indicators, one needs to take into account that:

-Indicators are a qualitative construct, not a scientific measurement; hence its interpretation requires utmost care in understanding its underlying theoretical/conceptual constructs and selection of data;

-Useful measurements are unique to each county; hence knowing the industrial structure of the country can clarify what are the information needed;

-More is not always better, all the elements need to be studied in the context and in proportion, coordination with other sector/activities and in sequence (order); (good interpretation requires to understand the context in which the indicator is used (be it a country, industrial structure or sector);

-No one prescription fits all, identify clear purpose of use; indicators are products of difficult compromises and one needs to know what has been compromised;

-Indicators are not written on stone, it will change with the changing reality; hence constant discussion, amendments and updates are expected. This is clear from series of revisions that has taken place in e.g. the Frascati and Oslo manuals.

REFERENCES

- Aboal, D., Mondelli, M.P., Vairo, M., (2015) Innovation and productivity in Agricultural Firms: Evidence from a Farm-level innovation Survey, Mimeo.
- Adam, F.(2014) Measuring National Innovation Performance : The Innovation Union Scoreboard Revisited, Springer Briefs in Economics, Hidelberg, Springer.
- Baptista, B., Bernheim, R., Garce, A, Hernandez, E., (2009) Consulta a Tomadores de Decision en Politicas Publicas de Ciencia, Tecnologia e Innovacion sobre sus Fuentes de Informacion: Informe Regional, Montevideo, Uruguay, ANII.
- Crespi, G. and Peirano, F. (2007) Measuring Innovation in Latin America: what we did, where we are and what we want to do, Paper prepared for the Conference on Micro Evidence on Innovation in Developing countries, Maastricht, the Netherlands, UNU-MERIT.
- De Beer, J., Fu, K., Wunsch-Vincent, S., (2013) The informal economy, innovation and intellectual property—Concepts, metrics and policy considerations. Economics Research Working Paper No. 10, Geneva, Switzerland, WIPO.
- Edquist, C and Zsbala-Iturriagoita, Jon Mikel (2015) The innovation Union Scoreboard is flawed: The case of Sweden-not the innovation leader of the EU-updated version, Papers in Innovation studies paper no. 2015/27, CIRCLE, Lund University.
- Foray, D and Hollanders, H (2015) An assessment of the Innovation Union Scoreboard as a tool to analyze national innovation capacities: The case of Switzerland, Research Evaluation, online version: doi: 10.1093/reseval/rvu036: pp. 1-16.
- Freeman, C., Soete, L., (2009) Developing science, technology and innovation indicators: What we can learn from the past, Research Policy, 38: 538-589.

- Freudenberg, M., (2003) Composite Indicators of Country Performance: a critical assessment, OECD Science, Technology and Industry Working papers, 2003/16, OECD.
- Gault, F. (2010) Innovation Strategies for a Global Economy: Development, Implementation, measurement and Management, Cheltenham, Edward Elgar.
- Gault, F (ed.) (2013) Handbook of Innovation Indicators and Measurement, Cheltenham, Edward Elgar.
- Grupp, H. and Moge, M.E. (2004) Indicators for national science and technology policy: how robust are composite indicators?, Research Policy, 33: 1373-1384.
- Hatzichronoglou, T. (1997), "Revision of the High-Technology Sector and Product Classification", OECD Science, Technology and Industry Working Papers, No. 1997/02. doi: 10.1787/134337307632.
- Hirsch Kreinsen, H., Schwinge, I., (2014) Knowledge-Intensive Entrepreneurship in Low-Tech Industries, Cheltenham, Edward Elgar.
- Huang, Can, Anthony Arundel & Hugo Hollanders, 2007, Non-R&D Innovation of Manufacturing Firms: Theory and Evidence from the Third European Community Innovation Survey, Innovation Watch. Brussels: European Commission.
- Iizuka, Michiko, Philippe Mawoko, & Fred Gault (eds.) 2015, Innovation for Development in Southern and Eastern Africa: Challenges for Promoting ST&I Policy UNU-MERIT POLICY BRIEF, 8pages.
- Intarakumnerd, Patarapong,(2007) Measuring innovation in Catching up Economies: An experience from Thailand. Paper obtained from internet accessed February, 2016.
- Konte, A. and Ndong, M.(2012) The Informal ICT sector and innovation Processes in Senegal, UNU Working paper 2012-009, Maastricht, UNU-MERIT.
- Mendonca, S. and N. von Tunzelmann (2004). Brave Old World: Accounting for "high tech" knowledge in "low tech" industries. DRUID Summer Conference 2004, Elsinore, Denmark, June 14-16.
- Mairesse, J., Mohnen, P., (2010) Using Innovations Surveys for Econometric Analysis, NBER Working Paper Series 15857, Cambridge, USA, NBER.
- Nardo, M. Saisana, M., (2008) OECD/JRC Handbook on constructing composite indicators. Putting theory into practice.
- OECD/ JRC (2008) Handbook on Constructing Composite Indicators: methodology and user guide, Paris, OECD
- OECD (2012) Recommendation of the Council on Regulatory Policy and Governance, Paris, OECD.
-

- OECD (2011) ISIC REV. 3 Technology Intensity Definition: Classification of manufacturing industries into categories based on R&D intensities, OECD Directorate for Science, Technology and Industry Economic Analysis and Statistics Division, Paris, OECD.
- von Tunzelmann, N. and V. Acha (2005), Innovation in 'low-tech' industries. Jan Fagerberg, David Mowery, Richard Nelson (eds), The Oxford Handbook of Innovation, New York: Oxford University Press: 407-432.
- Saltelli, A., (2007) Composite Indicators Between Analysis and Advocacy, Social Indicators Research, 81: 65-77.
- Saisana, M. Saltelli, A. and Tarantola, S. (2005) Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators, J.R. Statist. Soc.A 168 (2): 307-323.
- Schibany, A. and Streicher, G. (2008) How not to compare innovation performance: A critical assessment of European Innovation Scoreboard, Institute of Technology and Regional Policy-InTeREg
- Smith, K., (2005) Measuring Innovation, In Fagerberg, J., Mowery, D., Nelson R.R., Edited, The Oxford Handbook of Innovation, Oxford: Oxford University Press.
- Sutz, J (2012) Measuring innovation in developing countries: some suggestions to achieve more accurate and useful indicators, International Journal of Technological Learning Innovation and Development, 5 (1/2): 40-57.
- UNU-INTECH (2005) Measuring Innovation: Making innovation surveys work for developing countries, Technology Policy Briefs, 4 (1) 11 pages.



Analysing innovation policy indicators through a functional approach: the aeronautic industry case¹

Carolina Resende Haddad^{*}, Mauricio Uriona Maldonado^{**}

^{*}carol.resende.haddad@gmail.com

Department of Industrial and Systems Engineering, Federal University of Santa Catarina, Campus Universitario - Trindade, Florianopolis, 88040-270 (Brazil)

^{**}m.uriona@ufsc.br

Department of Industrial and Systems Engineering, Federal University of Santa Catarina, Campus Universitario - Trindade, Florianopolis, 88040-270 (Brazil)

ABSTRACT

Developing countries face different problems than developed countries and the use of the same indicator to evaluate and compare both regions can lead to misleading conclusions. Traditional indicators, such as R&D and patents may not capture the whole dynamic of a system, as they are used to compare systems focusing on its current structure. Many authors have been discussing the processes underlying industry transformation, innovation, and economic growth to access a system performance, i.e. the functions of innovation systems. Therefore, the purpose of this paper is to analyze these functions as indicators to measure the performance of the system in order to identify policy issues. In order to do that, we analyze the case of the aeronautic sectoral system of innovation of a region in Brazil. The functional approach helped us to better capture the dynamic of the system, by not restricting our analysis to the system's structure.

INTRODUCTION

In order to develop public policies to stimulate innovation towards local, regional or national needs, we need to understand how innovative a system is (Grupp & Schubert, 2010), which presupposes the ability to measure innovation. Therefore, many policymakers have discussed the development of indicators to better capture innovation activities (Gaut, 2013; Lee, 2015). Most parts of the typical innovation measurement tools are based on the linear model of innovation, i.e., on one or two indicators, such as patents and R&D (Mahroum & Al-Saleh, 2013). However, as argued by Archibugi, Denni and Filippetti (2009), many innovation indicators are not helpful for measuring innovation, as they do not reflect the innovation factors that distinguish different countries.

Lepori, Barré and Filliatreau (2008) pointed out that in the past decades there was the increase and diversification of STI indicators and innovation measurement in terms of analysis, types, consumers, and users, mainly due to the increasing complexity of the systems. For example, we can mention the use of composite innovation indicators (Grupp & Schubert, 2010). More recently, Mahroum and Al-Saleh (2013) proposed a measurement tool called the "Innovation Efficacy Index", which considers five functions of the "innovation through adoption" process

¹ This work was supported by the Industry Federation of Santa Catarina (FIESC).

This work is licensed under a Creative Commons License: Attribution-NonCommercial-NoDerivatives 4.0 International.

(accessing, anchoring, diffusing, creating and exploiting innovations) in order to understand cross-countries differences in innovation performance.

In this functionality line, many authors have been discussing the processes underlying industry transformation, innovation, and economic growth in order to evaluate the dynamic of the innovation system (Jacobsson, & Johnson, 2000; Liu, & White, 2001; Johnson, 2001; Hekkert *et al.*, 2007; Bergek *et al.*, 2008). These processes were labelled functions of innovation systems (Bergek *et al.*, 2008), which are the activities that take place in this system in order to generate technological change and disseminate innovations (Hekkert *et al.*, 2007). In this sense, the purpose of this paper is to analyze the functions of the innovation system as indicators to measure the performance of an emerging system of innovation of a developing country in order to identify policy issues.

We analyse the case of the aeronautic sectoral system of innovation in Santa Catarina State, Brazil. As pointed out by Hekkert *et al.* (2007), the functional approach (i) allows the comparison between innovation systems with different backgrounds; (ii) allows a systematic method of mapping the determinants of innovation; and (iii) allows the formulation of a set of policies that should be the target of the innovation system and the tools to achieve this target.

FUNCTIONS OF THE INNOVATION SYSTEM

The functional approach are related to the character and the interaction between the components of the innovation system (agents, networks and institutions) (Hekkert, & Negro, 2009). It was originally developed for Technology Innovation Systems, focused mainly on renewable energies (Jacobsson & Bergek, 2011; Negro, Hekkert & Smits, 2007). Other works extrapolated the renewable energy TIS and analyzed other IS, e.g. the ceramic tile Sectoral Innovation System (Gabaldón-Estevan; Hekkert, 2013).

Within the many attempts to identify functions, we will use the functions proposed by Hekkert *et al.* (2007), which are:

- Entrepreneurial activities: new entrants that identify an opportunity in the market and companies that diversify their business strategies;
- Knowledge development: mechanisms of learning, encompassing “learning by searching” and “learning by doing”;
- Knowledge diffusion through networks: is the exchange of information between actors in the innovation system;
- Guidance of the search: choose the focus of investments in technology among the options;
- Market formation: is the creation of protected spaces for new technologies, such as the formation of niche markets or by creating favourable tax regimes;
- Resources mobilization: is the allocation of resources, both financial and human capital, for specific technologies;
- Creation of legitimacy/counteracts resistance to change.

Table 1 shows the typical indicators to measure each of the seven functions (Hekkert *et al.*, 2007).

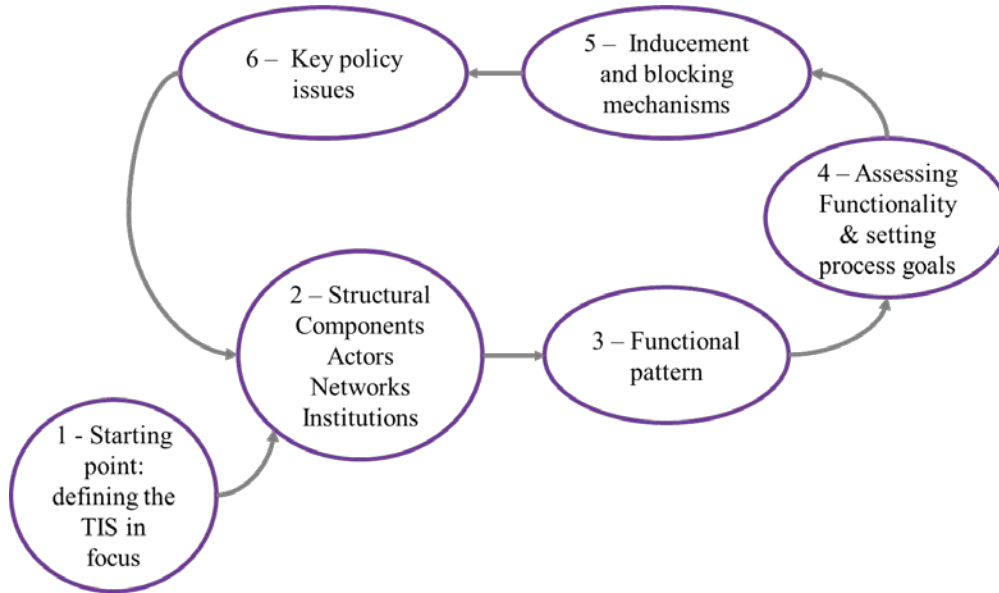
Table 1: typical indicators to measure the Functions of the Innovation System.

Function	Typical indicators
Entrepreneurial activities	Number of new entrants; Number of diversification activities; Number of new experiments with a new technology.
Knowledge development	R&D projects over time; Patents; Investments in R&D.
Knowledge diffusion through networks	Number of workshops and conferences on a particular technology The network size and intensity over time.
Guidance of the search	Specific targets set by governments or industries regarding the use of a specific technology; Number of articles in professional journals that raise expectations about new technological development.
Market formation	Number of niche markets that have been introduced; Specific tax regimes for new technologies; New environmental standards that improve the chances for new environmental technologies.
Resources mobilization	Funds made available for long-term R&D programs set up by industry or government to develop specific technological knowledge; Funds made available to allow testing of new technologies in niche experiments; Perception of the actors regarding the access to sufficient resources.
Creation of legitimacy/counteracts resistance to change	Rise and growth of interest groups; Lobby actions.

Source: adapted from Hekkert *et al.* (2007).

In order to analyse the SI and help policy makers in the selection and prioritization of public policies, Bergek *et al.* (2008) proposed an analytical scheme that allows accessing the performance of the system and identifying the aspects that are influencing this performance:

Figure 1: Scheme of analysis.



Source: adapted from Bergek *et al.* (2008).

METHOD

Based on the main aim of our paper, we followed the six steps proposed by Bergek *et al.* (2008) to analyse the functional dynamics of the aeronautic industry in Santa Catarina's State:

- Step 1: we defined the focus of the Sectoral System of Innovation;
- Step 2: we identified the structural components of the innovation system;
- Step 3: we mapped the functional pattern of the system considering the seven functions of the innovation system proposed by Hekkert *et al.* (2007). In this step, we collected the data for document analysis and expert's interviews;
- Step 4: we accessed the functionality of the system by analysing its phase of development and set the final process goal;
- Step 5: we identified the system's inducing and blocking mechanism;
- Step 6: we specified the key policy issues concerning the aeronautic industry final process goal.

We conducted expert panels in order to collect data to steps 4 to 6.

THE AERONAUTIC INDUSTRY CASE

The system that will be the focus of this paper is the aeronautic industry of Santa Catarina's State (SC), Brazil. A survey made by the Industry Federation of Santa Catarina (FIESC, 2013) identified that this Sectoral Innovation System was diagnosed to be in an emerging level.

Step 1 – Define the system's focus

To set the focus of the Innovation Sector Aeronautic system in SC, we used the Brazilian National Classification of Economic Activities (CNAE) number 30, subsections number 30.4, 30.5 e 30.9, which are specifically related to the aeronautic industry.

Step 2 – Define the structural components

Due to the emerging level of the industry analysis, it was difficult to clearly identify those actors, networks and institutions that strongly influence the aeronautic industry. Thus, we will discuss in general terms those who are nowadays are present at this stage of the industry.

In term of actors, SC is characterized by the presence of few companies in the sector. FIESC is an active actor in the industry, as it represent all the industries in the state. We can also mention as actors, labour unions and regulatory agencies. Some universities and regional research institutes act as actors in the system, but they have no prominent role in the aeronautic system yet.

Considering the aeronautic system networks, we identified the relationship between suppliers of different levels of the supply chain and between companies and labour unions. The university-enterprise network little influences the system nowadays.

We observed a lack of institutions of interest in the system, especially because it is still in an emerging stage. The National Civil Aviation Agency (ANAC) is one of the agencies that regulate the sector in Brazil.

Step 3 – Analyse the functionality

We describe the main findings of the aeronautic system functionality below.

Entrepreneurial activities

Number of new entrants

- It is expected that Novaer, a new entrant that produces small aircraft, will install a factory in Lages, in the South of SC;
- Lack of entrepreneurs and companies specialized in the production of high value-added products for the aeronautic industry.

Number of diversification activities

- There is little diversification activities in the aeronautic industry.

Number of new experiments with a new technology

- New experiments are being developed by entrepreneurs in small aircraft;
- Lack of experiments in fuselage.

Knowledge development

R&D projects over time

- Knowledge is being developed through “learning by doing” in the construction of small aircraft;

- Researches are being developed in kits for light aircraft;

Patents

- A search in the Brazilian National Institute of Industrial Property (INPI), we could find 49 patents registered with “aeronautic” in the title (INPI, 2016).

Investments in R&D

- Novaer Craft (2015) plans to install an engineering center in Florianopolis in order to develop research in the aeronautical sector.

Knowledge diffusion through networks

Number of workshops and conferences on a particular technology

- The Development Committee of the Aeronautical Industry is responsible for the diffusion of knowledge in the sector by promoting courses focused on the industry (FIESC, 2016);
- Some actors in the IS are participating in international conferences and workshops, such as the SUN’n FUN in the United States, which help to approximate regional companies to international ones.

The network size and intensity over time

- Lack of interaction between the university, government and industries, which could be better explored to diffuse knowledge through the system networks.
- The *Brazilian Aerospace Cluster* Project, developed by the Brazilian Spatial Agency (AEB, 2014), aims at creating an aeronautic industry cluster in SC.

Guidance of the search

Specific targets set by governments or industries regarding the use of a specific technology

- The Development Program for Certification to Small-Sized Aircraft (iBR2020), prepared by ANAC (2015a), develops projects focused on small aircraft in order to make them more prepared to succeed when subjected to certification;
- Another factor that influences the direction of search are the standards for the aerospace industry based on ISO 9001, from the AS/EN 9100 series;
- Other certifications related to the sector are ISO 14.001 (environmental management), Management System of Occupational Health and Safety - OHSAS 18.001 (2015), the Type-certificate for aircraft, engines and propellers, and the Aircraft Certificate in transport category (ANAC, 2015b);
- The *Federal Aviation Administration* (FAA), from the EUA, and the *European Aviation Safety Agency* (EASA) regulate the civil aviation and influence the guidance of the search;
- The Development Committee of the Aeronautical Industry influences the direction of search since it promotes discussions about the industry's technological guidelines in the state.

Number of articles in professional journals that raise expectations about new technological development

- We did not find any aspect related to this indicator.

Market formation

Number of niche markets that have been introduced

- We did not find any niche market that has been introduced;

Specific tax regimes for new technologies

- There are no tax regimes to encourage the development technologies such as unmanned aerial vehicle (UAV), transmission line, software, and aeronautical instruments for agriculture;
- The Plano Brasil Maior (2014) is a program that aims to enhance and build new technological competencies by promoting tax incentives and can influence market formation;
- The Inova@SC Program seeks to promote the development of innovation in SC and can influence market formation;

New environmental standards that improve the chances for new environmental technologies

- New environmental standards and legislations, such as the Greenhouse Gas (GHG, 2016) Protocol, can influence market formation for the development of new technologies that reduce environmental impacts.

Resources Mobilization

Funds made available for long-term R&D programs set up by industry or government to develop specific technological knowledge

- Educational initiatives for resources mobilization in the aeronautic industry with the creation of technological courses, such as Aerospace Engineering from UFSC and the Technical Course in Aircraft Maintenance (for avionics, cell and propellant engines) from SENAI;
- Another program set by the government is the iBR2020, created to develop education initiatives in order to improve the national aircraft industry capacity to develop small aircraft designs that are more able to succeed when subjected to the Type-certificate.

Funds made available to allow testing of new technologies in niche experiments

- The “Plano Brasil Maior” (2014) is a source of financing new technologies and can influence resources mobilization in the aeronautic industry.

Perception of the actors regarding the access to sufficient resources

- Lack of professionals specialized in services for the aeronautic industry.

Creation of legitimacy/counteracts resistance to change

Rise and growth of interest groups

- Lack of activities related to the creation of legitimacy in the aeronautic sector in SC;
- FIESC is able to bring legitimacy to the sector and has been considered the articulator of the voice of the aeronautic industry nowadays

Lobby actions

- Lobby to build a new Aviation School at SENAI (FIESC, 2013a);
- Lobby to expand the aeronautic network in SC and bring new enterprises.

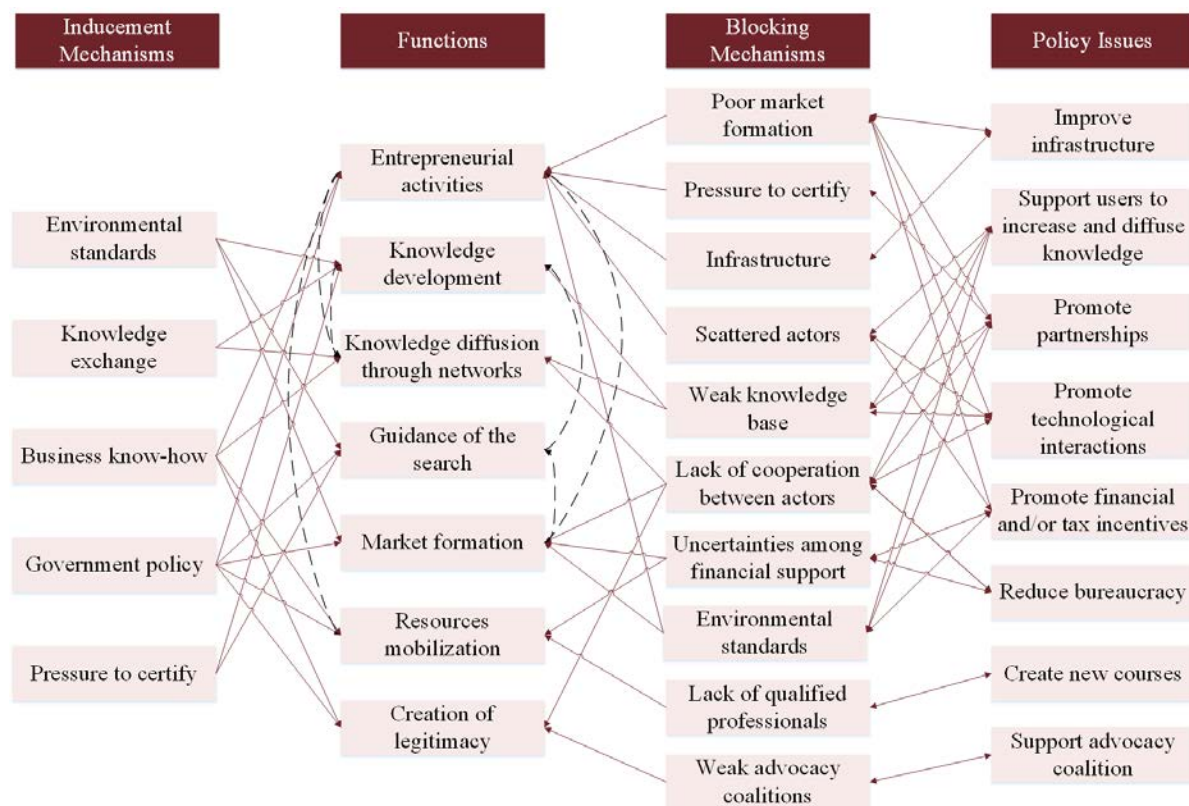
Step 4 – Assessing functionality and setting process goals

After analysing the system's functionality, we identified, as previously acknowledged by FIESC (2013a), that the aeronautic industry is in a formation phase. Therefore, the final process goal is to develop an aeronautic sector recognized in the formation of specialized human resources, technology development and manufacturing of small aircraft.

Step 5 – Inducing and blocking mechanisms

Figure 2 shows the mechanisms that induce and block the development of the functions of the aeronautic industry. We identified five main inducing and ten main blocking mechanisms. In addition, the dotted arrows show the functions interdependencies.

Figure 2: Inducing and blocking mechanisms and key policy issues of the emerging aeronautic sector in SC State.



Source: Uriona & Haddad (2015).

Step 6 – Key policy issues

From the analysis of the system functions and the inducement and blocking mechanisms, we obtained the following key policies (see the fourth column of Figure 2):

- Encourage the formation of partnerships in order to develop markets and induce the development of research and partnerships between actors of the systems;
- Provide tax and financial incentives to stimulate the development of markets and encourage new technologies and the development of researches;
- Improve infrastructure in order to stimulate market formation;
- Encourage research development and knowledge dissemination between the actors of the system;

- Promote the approach of the actors of the system in order to promote the development of studies on new technologies;
- Reduce bureaucratic procedures in order to encourage market formation;
- Create new courses to develop skilled labour and expand the number of courses offered in the aeronautic sector;
- Support group coalition in order to create groups of interest and encourage new technologies.

CONCLUSION

Out purpose in this paper was to use the functional approach to analysing the performance of an innovation system in order to identify key policy issues. By analysing the aeronautic industry in a region of a developing country, we could find out, by considering a number of indicators, how this sectoral innovation system is currently functioning. Then, we showed the inducing and blocking mechanisms and pointed out key policy issues.

The functional approach also enabled, in this case, to capture factors in the system that common innovation measures would not capture. For example, the function entrepreneurial activities showed new experiments that are being developed by entrepreneurs in small aircraft.

To sum up, the functional approach helped us to better capture the dynamic of the system, by not restricting our analysis to the system's structure. Therefore, we claim that the functional approach can better capture the dynamics of a specific country and better reflect the innovation factors that distinguish different countries or regions. As an opportunity for further research, we suggest the use of a broader range of indicators to analyse the system functioning, such as the ones that consider the specific characteristics of a developing country, as mentioned in the Oslo Manual.

REFERENCES

- Archibugi, D., Denni, M., & Filippetti, A. (2009). The technological capabilities of nations: The state of the art of synthetic indicators. *Technological Forecasting and Social Change*, 76(7), 917-931.
- Bergek, A., Jacobsson, S., Carlsson, B., Lindmark, S., & Rickne, A. (2008). Analyzing the functional dynamics of technological innovation systems: A scheme of analysis. *Research Policy*, 37(3), 407-429.
- FIESC. (2013a). FIESC specialized cameras activity report. Florianópolis: FIESC.
- FIESC. (2013b). Future sectors for Santa Catarina industry in 2022: a socioeconomic study. Florianópolis: FIESC.
- Gabaldón Estevan, D., & Hekkert, M. P. (2013). How Does the Innovation System in the Spanish Ceramic Tile Sector Function?. *Boletín de la Sociedad Española de Cerámica y Vidrio*, 52(3), 51-158.
- GHG. *Greenhouse Gas Protocol*. GHG Protocol Web site. <http://www.ghgprotocol.org/>, accessed March 2016.

Grupp, H., & Schubert, T. (2010). Review and new evidence on composite innovation indicators for evaluating national performance. *Research Policy*, 39(1), 67-78.

Hekkert, M. P., & Negro, S. O. (2009). Functions of innovation systems as a framework to understand sustainable technological change: Empirical evidence for earlier claims. *Technological Forecasting and Social Change*, 76(4), 584-594.

Hekkert, M. P., Suurs, R. A., Negro, S. O., Kuhlmann, S., & Smits, R. E. (2007). Functions of innovation systems: A new approach for analysing technological change. *Technological Forecasting and Social Change*, 74(4), 413-432.

INPI. (2016). *INPI patent base*. INPI Web site. <http://www.inpi.gov.br/menu-servicos/informacao/bases-de-patentes-online>, accessed March 2016.

Jacobsson, S., & Bergek, A. (2011). Innovation system analyses and sustainability transitions: Contributions and suggestions for research. *Environmental Innovation and Societal Transitions*, 1(1), 41-57.

Jacobsson, S., & Johnson, A. (2000). The diffusion of renewable energy technology: an analytical framework and key issues for research. *Energy Policy*, 28(9), 625-640.

Johnson, A. (2001). Functions in innovation system approaches. *Proceedings of the DRUID's Nelson-Winter Conference*, Copenhagen, Denmark.

Lee, Y. N. (2015). "Evaluating and extending innovation indicators for innovation policy", *Research Evaluation*, 24, pp. 471-488.

Lepori, B., Barré, R., & Filliatreau, G. (2008). New perspectives and challenges for the design and production of S&T indicators. *Research Evaluation*, 17(1), 33-44.

Liu, X., & White, S. (2001). Comparing innovation systems: a framework and application to China's transitional context. *Research Policy*, 30(7), 1091-1114.

Mahroum, S., & Al-Saleh, Y. (2013). Towards a functional framework for measuring national innovation efficacy. *Technovation*, 33(10), 320-332.

Negro, S. O., Hekkert, M. P., & Smits, R. E. (2007). Explaining the failure of the Dutch innovation system for biomass digestion—a functional analysis. *Energy Policy*, 35(2), 925-938.

OECD (2005). *Oslo Manual*. Paris: OECD.

_____. (2010). *Ministerial report on the OECD innovation strategy: innovation to strengthen growth and address global and social challenges*. OECD Web Site <<https://www.oecd.org/sti/45326349.pdf>>, accessed March 2016.

Uriona, M. M. & Haddad, C. R. (in review). A proposed innovation system foresight method for developing countries. *Technological Forecasting and Social Change*.



Assessing the performance of national innovation systems in Europe

Charles Edquist*, and Jon Mikel Zabala-Iturriagoitia**

* charles.edquist@circle.lu.se

CIRCLE, Lund University, Solvegatan 16 PO Box 188, Lund, 22100 (Sweden)

** jmzabala@deusto.es

Deusto Business School, University of Deusto, Camino de Mundaiz 50, Donostia-San Sebastian, 20012 (Spain)

ABSTRACT

To support the establishment of the European Innovation Union, the European Commission is using the Innovation Union Scoreboard (IUS). In this paper, the performance of EU28 national innovation systems are analyzed from an efficiency perspective by using exactly the same data as those provided by the IUS for years 2010-2015. This efficiency analysis was carried out using Data Envelopment Analysis.

Our analysis demonstrates that the results based on efficiency measures reflect that in general terms innovation systems are widely underexploited in Europe and that there are important variances among territories. We have shown that many countries which devote fewer resources than the innovation leaders, achieve outstanding levels of efficiency and, contrary to what the IUS predicts, countries with consolidated innovation systems, do not show efficiency levels commensurate with their expected competitiveness.

INTRODUCTION

The European Commission highlights that, regarding science and innovation, Europe is one of the innovation intensive regions in the world. With the recent strategy “Europe 2020”, Europe is focusing on today’s challenges in a changing world and wants to become a sustainable, inclusive and smart economy. Thereby the European Union has set ambitious objectives in five areas to be reached by 2020. Besides climate and energy, education, employment and social inclusion, innovation is one of these five pillars to form a so called “European Innovation Union”.

To support the establishment of the European Innovation Union, the European Commission is using the Innovation Union Scoreboard (IUS) as a tool to monitor the implementation and to examine the innovation performance of European member states and evaluate (and rank) their research and innovation systems. This means that the IUS is intended to have a real impact on the evaluation of the policies of the Member States, the allocation of resources and – supposedly – for the design of innovation policies at the European, national and regional levels. Hence, the design of the IUS and its results are supposed to have real impact.

To assess the innovation performance of the member states, a Summary Innovation Index (SII) is provided by the IUS. In 2014, the SII includes 25 indicators, which are equally weighted. These indicators are divided into three main categories (i.e. enablers, firm activities and outputs) and eight dimensions (i.e. human resources, excellent research systems, finance and support, firm investments, linkages and entrepreneurship, intellectual assets, innovators, economic effects). According to this single synthetic composite indicator Denmark, Finland, Germany and Sweden are the innovation leaders within the EU.

As to the SII, for each year, each indicator has a normalized score that varies from a minimum performance of 0 up to maximum of 1. The normalized scores are added to each other and divided by the number of dimensions within each of the eight dimensions. Afterwards the composite index (SII) is formed by calculating the average among all eight dimensions. Countries are ranked according to the SII in the following groups: innovation leaders (more than 20% above EU average), innovation followers (less than 20% above, or more than 90% of the EU average), moderate innovators (relative performance rates between 50% and 90% of the EU average) and modest innovators (less than 50% of the EU average).

The IUS measures the innovation performance for each country by summarizing all 25 indicators into a single SII, irrespective of whether the indicators are presenting innovation outputs or innovation inputs – or something else. We argue that synthetic or composite innovation measures such as the one provided by the IUS (i.e. SII) are highly misleading. The data (the separate indicators) need to be analyzed much more in depth in order to reach comprehensive views of the performance of an innovation system. In this paper we compare the input and output indicators of the IUS (as is done in productivity and efficiency measurements), what provides a measure of innovation performance that complements the information provided by the synthetic indicator of the IUS.

EFFICIENCY CONSIDERATION

To measure performance/productivity/efficiency of innovation systems, the indicators need to, in some way, be split up into innovation inputs and clear innovation outputs. Both sides need to be considered separately, and then related to each other. Only in this way can the efficiency or productivity of national systems of innovation be estimated and compared. Neither input nor output indicators themselves can measure the innovation performance of a country. It is rather the relation and balance between the input and output side which measures how a country is performing in its innovative actions. To be able to assess which of the indicators provided by the IUS are input and output indicators respectively, we define inputs and outputs as follows:

- Innovation input indicators refer to the resources (human, material and financial; private as well as governmental) which are used to create innovations, including bringing them to the market.
- Innovation output indicators refer to new products and processes, new designs and community trademarks as well as marketing and organizational innovations, which are new to the market and/or new to the firm and are adopted by users.

In this paper, the performance of EU28 national innovation systems are analyzed from this efficiency perspective by using exactly the same data as those provided by the IUS. We start by gathering all the data from the IUS reports between 2010 and 2015.¹ These data are all normalized scores for each indicator and for all EU28 countries. We then rank all countries, as well as the EU28 average, for each indicator. This gives us the opportunity to make a preliminary analysis of the relative position of each national innovation system in a diverse set of measures.

METHODOLOGY

The innovation performance in efficiency terms is measured as the relation between the inputs and the outputs. By grouping the IUS indicators in inputs and outputs, we are able to see the extent at which innovation inputs are transformed into or materialize into innovation outputs. From our point of view it is important to see the relationship between the input and the output side and assess their balance. A high level of the input indicators means that large efforts and resources are devoted to stimulate innovation. Similarly, a high score for the output shows

that a country has a high production of innovations. However, if the input side is, relatively speaking, much larger than the output side, the efficiency of the system as a whole is low. This efficiency analysis was carried out using Data Envelopment Analysis (DEA). Any estimated efficiency score refers to the spatial performance of a particular decision making unit (i.e. the decision making units in our case are national innovation systems), and, thus, can be used to assess the performance of the entire system, by establishing a fictitious optimum or benchmark by linear combination of the most efficient units given the ratio of their outputs to their inputs and relating observations to that level. From this point of view, innovation systems are depicted as technically more or less efficient transformers of inputs into outputs. There are two general approaches to measuring efficiency: (1) parametric models, such as Stochastic Frontier Analysis, and (2) non-parametric models, such as Data Envelopment Analysis (DEA) and Free Disposal Hull. In this paper we will use DEA for this efficiency analysis. DEA takes a systems approach, which takes account of the relationship between all inputs and outputs simultaneously, without requiring a weighting system that reduces these units into a single unit measure, as each input or output can be measured in its natural physical units. In addition, DEA does not impose any preconceived functional form on the data when determining efficient units, so the production function of efficient units is estimated using piecewise linear programming on the sample data rather than making restrictive assumptions about the underlying production technology. The importance of this feature is that a unit's efficiency can be assessed based on the performance observed by others (i.e. or any linear combination derived thereof). In turn, DEA identifies the inefficiency in a particular unit by comparing it to similar units regarded as being efficient, rather than by trying to associate a unit's performance with statistical averages. The principal disadvantage of DEA, is that it assumes the number of units and data included to be free of measurement errors. DEA is thus particularly sensitive to unreliable data because the efficient units determine the efficient frontier and, thus, the efficiency values of those units behind this frontier. Thus, the number of efficient units at the frontier tends to increase with the number of inputs and output variables, which results in loss of discriminatory power of this method. By using the official data published by the IUS we expect that this disadvantage is minimized as much as possible. Figure 1 depicts the general idea of the frontier concept used in DEA.

Figure 1: Frontier concept and efficiency calculation.

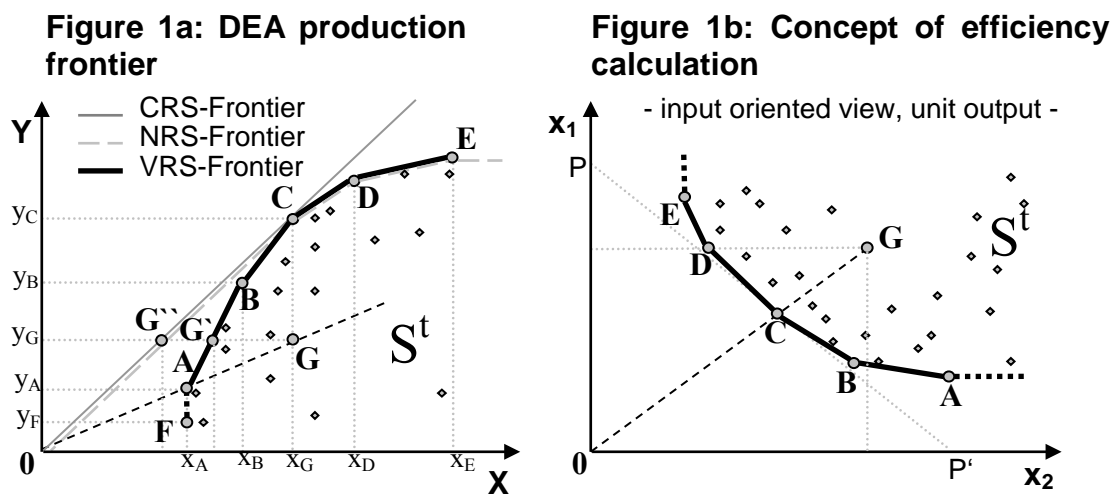


Figure 1a depicts a production frontier (isoquant) by means of an XY-coordinate system whereas points A, ..., E define the scope and shape of the frontier, S^t refers to the production possibility set in time t , and CRS, NRS, and VRS are frontiers with Constant Returns to Scale [RtS], Non-increasing RtS, and Variable RtS, respectively. Points F and G lie below the frontier and illustrate inefficient input/output combinations. The technical efficiency [TE] of point G can be obtained by calculating $(X_G)/(X_G)$. The calculation of this measure can be illustrated even better in a two-dimensional X_1X_2 -frame (two inputs applied to produce one unit output) as in Figure 1b. Points A, ..., E all refer, once again, to (technically) efficient combinations of X_1 and X_2 in order to produce one unit of output and they therefore define the frontier. Point G corresponds to an inefficient observation since X_1 and X_2 can be reduced without any drop in output. The TE of G can be obtained by calculating $\overline{OC}/\overline{OG}$. Hence, TE has a range $0 \leq TE \leq 1$, whereas 1.0 refers to a best practise, fully efficient example. Concerning point G, however, one has to reduce both inputs, e.g. in the proportion $1-TE$, in order to be efficient (reach the frontier).

We estimate the production set S^t and the corresponding frontier by considering:

$$S_{DEA}^t = \left\{ (x, y) \in R_+^{p+q} \mid y \leq \sum_{i=1}^n \gamma_i y_i ; \quad x \geq \sum_{i=1}^n \gamma_i x_i \quad \text{for } (\gamma_1, \dots, \gamma_n) \right\}$$

such that $\sum_{i=1}^n \gamma_i = 1; \gamma_i \geq 0, i = 1, \dots, n; x$ can produce y with $x \in R_+^p$ and $y \in R_+^q$

which refers to the smallest free disposal convex set covering all the data.

RESULTS

The results obtained with this approach helped us reach an overall assessment of EU28 national innovation systems and their innovation performance. As our results evidence, this efficiency approach shows a strong degree of complementarity with those provided by the IUS. The rationale for using efficiency to complement the conclusions obtained through the IUS lies in the fact that the latter follows a “the more the better” logic. Namely, the more resources (inputs) a country puts into the system, the more competitive it will be – more innovations (outputs) obtained. The efficiency measurement approach aims at providing information about the use (misuse) of these resources. The efficiency of resource use is indicated by the degree to which these inputs generate soaring returns, or output results that do not reflect the level of investment.

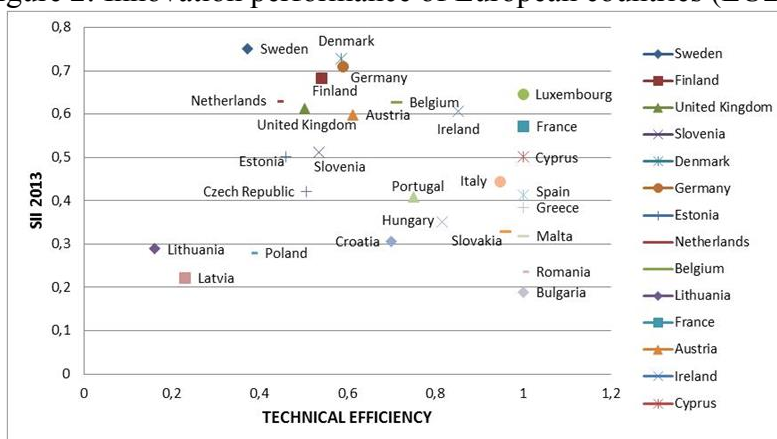
The overall mean of the calculated TE for the EU28 countries studied was 0.702 in year 2013 (std. 0.265 and typical error 0.05). Even if this value should be regarded as positive, it also indicates that there is a still potential for improving innovation performance in Europe. In other words, according to our empirical results, innovation potential is still widely underexploited in Europe (by almost 30% on average).

Our results reveal that a number of countries (i.e. eight in particular) had highly efficient innovation systems (see Table 1 and Figure 2). In fact, this is the case for France, Cyprus, Luxembourg, Spain, Greece, Romania, Malta and Bulgaria. Theoretically, most observations (i.e. countries) could be expected to be close to the frontier, and to behave as efficient units, but as the table below shows, there is wide variance in innovation performance in Europe.

Table 1. Technical efficiency of European countries (EU28).

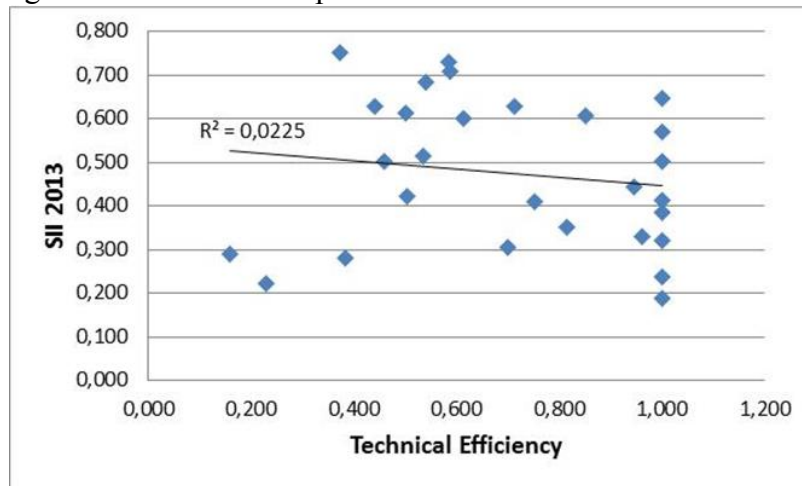
	CRS	VRS	Scale efficiency	SII 2013
Sweden	0,372	0,743	0,501	0,750
Finland	0,540	1,000	0,540	0,684
United Kingdom	0,502	0,503	0,998	0,613
Slovenia	0,534	0,580	0,921	0,513
Denmark	0,585	1,000	0,585	0,728
Germany	0,589	1,000	0,589	0,709
Estonia	0,459	0,757	0,605	0,502
Netherlands	0,441	0,772	0,571	0,629
Belgium	0,713	0,989	0,720	0,627
Lithuania	0,160	0,181	0,887	0,289
France	1,000	1,000	1,000	0,571
Austria	0,613	1,000	0,613	0,599
Ireland	0,853	0,902	0,946	0,606
Cyprus	1,000	1,000	1,000	0,501
Luxembourg	1,000	1,000	1,000	0,646
Czech Republic	0,505	0,706	0,716	0,422
Poland	0,383	0,466	0,821	0,279
Slovakia	0,960	1,000	0,960	0,328
Croatia	0,699	0,761	0,918	0,306
Latvia	0,229	0,300	0,763	0,221
Portugal	0,751	0,999	0,752	0,410
Hungary	0,814	0,873	0,933	0,351
Spain	1,000	1,000	1,000	0,414
Italy	0,947	1,000	0,947	0,443
Greece	1,000	1,000	1,000	0,384
Romania	1,000	1,000	1,000	0,237
Malta	1,000	1,000	1,000	0,319
Bulgaria	1,000	1,000	1,000	0,188

Figure 2: Innovation performance of European countries (EU28).



With regard to the position of each country in relation to the frontier (level, near, far away) and its related TE score, all observations can be ordered by their achieved innovation efficiency. This ranking was compared with that provided by the SII, which according to the IUS, measures innovation competitiveness of European countries. In Figure 3 the two rankings are related: the y-axis refers to the SII index and the x-axis to the efficiency based values (TE).

Figure 3: The relationship between SII and Technical efficiency.



If the two performance indicators coincided, one would expect the majority of points to be along a 45° line. However, as Figure 4 shows, this is not the case. Indeed, the trend line indicates a negative relationship. To some extent, the rankings are reversed; therefore, as argued above, radically changing the 'blueprint' on which policy recommendations are based. The negative relation of these indices must result from their different conceptual settings, since the measures employed in both cases are the same. While the SII is created as a measure mainly oriented to the inputs in the system in the sense of 'the more the better', the efficiency measure refers to the how these resources are used relative to a particular output. In addition, the efficiency index allows a comparison between the difference levels of innovation performance since it compares among countries.

Thus, although a country that is at the top of the TE ranking, but which employs very few resources might be efficient in terms of resource use (top in terms of TE), in terms of enhancing regional development, closing the gap in growth rates, social welfare, etc. this same country might be contributing very little and be classed as lagging. On the other hand, a country that invests huge amounts of resources to improve its innovation system (i.e. is top in terms of SII), but whose use of resources is identified as inefficient compared to the peer group of best practice regions, cannot be seen as an example of best practice. Hence, in order to assess the performance and institutional quality of an innovation system, both aspects must be considered.

DISCUSSION

The IUS identifies those countries with high investment in high-tech related activities as the leading countries. Our analysis demonstrates that the results based on efficiency measures reflect that in general terms innovation systems are widely underexploited in Europe and that there are important variances among territories. We have shown that many countries which devote fewer resources than the innovation leaders, achieve outstanding levels of efficiency

and, contrary to what the IUS predicts, countries with consolidated innovation systems, do not show efficiency levels commensurate with their expected competitiveness.

Even if the “innovation leaders” of the EU may be regarded as comprehensive in many aspects, the results indicate that their efficiency levels are far from being adequate. The innovation leaders, generally speaking, invest vast resources and still do not manage to produce as much outputs as other countries. The results we obtained might perhaps be explained by the complexity of innovation processes and thus the need to coordinate the activities promoted by innovation policies. Those countries with higher R&D expenditure levels, and which have a long tradition in the implementation of science, technology and innovation policies, tend to support new growth industries which imply higher risks in their innovation policy proposals. As a result, the innovation systems in these countries devote more inputs, which despite render the systems very dynamic, the high levels of coordination required and the uncertainties involved reduce their levels of efficiency. Similarly, those territories with lower absorptive capacity and fewer resources, adopt the embodied knowledge and the innovations of others, which involve lower levels of development and at the same time produce more efficient behaviors since risk is avoided and the 'new' knowledge is rapidly adopted. It also to note that the countries with fewer resources to invest have to pay much more attention to how they are used. They cannot afford to squander the scarce resources dedicated to innovation activities. Their more cautious behavior produces unexpected and unforeseen efficiencies.

From a quantitative perspective, the approach followed by the IUS seems to offer a partial view of the actual state of innovation systems. We have shown that the use of these indicators within different methodological frameworks yields differing, but not necessarily contradictory results. Thus, they provide a partial picture of the phenomenon being examined; different approaches should be seen as being complementary. Therefore, policy makers will need to consider the results of different and complementary analyses to obtain a comprehensive picture of their respective innovation system. From our point of view, the sum of each partial view will provide a clearer picture than that provided by each in isolation.

ⁱ In this abstract, and due to length constrains, only the preliminary results for the IUS 2013 are reported.



The process of construction of evidence: An analysis of the use of indicators in two decisions of innovation policy

Nuno Boavida*, António Moniz* and Manuel Laranja**

* nuno.boavida@campus.fct.unl.pt

CICS Nova of Universidade Nova de Lisboa.

** ISEG of Universidade de Lisboa.

ABSTRACT

Despite increasing calls for evidence-based policies, knowledge about the practical use of evidences remains limited. This paper studies the process of construction of evidences in decisions of innovation policy to understand how evidences were used. It analysis the use of indicators and other evidences through interviews conducted to inquire about the two decisions: an electric mobility policy and a nanotechnology laboratory. Results show indicators and other evidences were brought to decision processes according to their availability and capacity to support the different interests of the actors and the stakeholders. Their role was influenced by the particular situation of the decision makers. More importantly, the use of persuasive analytical evidences appears to be related with the adversity of the policy context. In addition, research suggests that indicators are one tool among others to foster innovation decisions. In fact, the relatively minor instrumental role of indicators suggests that indicators are mostly a complementary instrument of decision. When used relevantly, indicators can offer support to a decision. But there are other significant influences that need to be taken into account to understand the specific role indicators and other evidences play, such as the social relations of the decision makers and their emotional-intuitive decisions.

KEYWORDS

Evidences; indicators; innovation policy; decisions.

1 THE USE OF EVIDENCE

The use of evidences in policy-making has seen a growing interest in recent years. Several researchers reported an increase in calls for public policies that use evidences.¹ The calls assume that evidence-based policies are an aspiration that improves the quality of the decision, , among other things (C. Porter 2010). However, despite these calls, knowledge about the use of evidence in practice remains rather limited. In fact, we know very little about policy makers' use of information in practice, how information is valued and, in particular, what is the prevalence of formal scientific evidence use in policy decision (Hall and Jennings 2010). Hence, there is the need to study the practices of use of evidence to be able to support claims for evidence-based policies. A way to understand the use of evidences in practice is to centre the study in the use of indicators in cases of policy making. Indicators are conceptual instruments used to measure, evaluate and help with decisions by summarizing characteristics or highlighting what is happening in reality. They are closer to formal scientific evidences and frequently quantified. Furthermore, in a study about sustainable policy at the EU institutions, Sébastien and Bauler (2013) pointed to the need to study the processes of construction of evidence to understand the use of indicators, rather than to focus on their technical quality and their independency from their producers (two factors initially presumed important). Thus, there is the need to study the process of construction of evidences in policy decisions to understand how indicators and other types of evidences are used in practice.

There are several factors that can account for our limitation of knowledge regarding the use of evidences. First, the novelty of the field naturally limits information about the use of evidence in practice. Second, there is an abundance of definitions of evidences that limit our ability to report their use. In one extreme, evidences can be strictly identified with scientific outputs. In this case, evidences comprises all types of science (and social science) knowledge generated by a process of research and analysis, either within or without the policy-making institution (Juntti, Russel, and Turnpenny 2009). On the other end of the spectrum, evidences are interpreted as pieces useful to support policy. In this case, an evidence is not necessarily data or information, but mostly a selection of the available information introduced in an argument to persuade about the truthiness or falsity of a statement (Flitcroft et al. 2011). Third, evidences can assume various forms in different contexts which limit their identification. In fact, evidences can be indicators, historical facts, statistics², and results of experiments, texts, quotes from secondary sources, real experiences or histories, or opinions of individuals in one field. Fourth, these forms can vary with the context: In policy-making, evidences can range from numerical data to ethical/moral interpretations expressing values, attitudes and perceptions of stakeholders and other decision makers. In health contexts, evidences can be research findings, other knowledge that is explicit, systemic and replicable, or simple acceptable waiting times (Lomas et al. 2005). In management contexts, evidences can include costs, technical characteristics of materials, stakeholders' opinions, etc.

The use of evidences in policy-making can be a significant subjective process. In fact, the strength and quality of evidence can be related to the number of controversies that it goes through during its lifetime (Sébastien and Bauler 2013). In these cases, evidence loses strength in the process of decision-making with the increase of controversies it goes through since its creation. Furthermore, the selection of evidence can also depend on the situations in which policy makers find themselves. These situations shape which information is used from the complex set available, and which evidence is rejected or at least downplayed (Perri 6 2002). In fact, policy-making “always makes use of some evidence, but there is a plurality [...] of things that count as evidence, and what counts depends on where policy makers are situated” (Perri 6 2002, 7). In addition, the selection of evidences can be related to epistemological choices of the decision maker, in terms of claims about valid sources of knowledge and how to judge knowledge claims. These choices can be related to the use of quantitative or qualitative information, but also sometimes religious believers might endorse theological claims to knowledge. The choices often reflect ontological assumptions about the objectivity or subjectivity of reality. For example, for some only positivistic techniques of inquiry support claims to knowledge as

¹ See among others Head (2010), Flitcroft et al. (2011); Juntti et al. (2009), Sorrell (2007), and Hall & Jennings (2010).

² For the purpose of this work a statistic is a numerical fact or datum, especially one computed from a sample (Gault 2013).

reliable facts, whereas for others the complexities of the social world demand an interpretation of human behaviour and intentions (Henn et al. 2009). In this context, policy emerges from the interaction of different forms of evidences, filtered and shaped by the processes of decision-making (Flitcroft et al. 2011, 1039).³ These filtering processes are subjective and evidence can be chosen instead of another, leading the argument in different directions. Therefore, what counts as evidence in policy-making is subjective, depends on the controversies associated to them, the particular situation of the decision maker and their epistemological claims.

2 INDICATORS IN INNOVATION POLICY

Policy-making can use many types of evidences, as mentioned earlier. This work will place a particular emphasis on a concrete type of evidence closely related to scientific findings: indicators. There are many indicators that can be used, specifically, for innovation policy purposes. In fact, the last decades have witnessed a significant amount of literature and other efforts directed to developed indicators in this policy field. In this context, indicators can be defined as conceptual instruments used to measure, evaluate and help with decisions by summarizing characteristics or highlighting what is happening in reality. They are commonly understood as variables selected to characterize the efforts undertaken by countries/regions/companies in the field of science and technology and innovation. These indicators cover resources devoted to research and development (R&D), innovation, patenting, technology balance of payments, international trade in R&D-intensive industries, etc. There is a significant amount of innovation indicators freely available for comparisons over time and across countries, regions, sector and companies. For example, the Innovation Union Scoreboard captures the economic success of innovation using five indicators: 'Employment in knowledge-intensive activities', the 'Contribution of medium and high-tech product exports to the trade balance', 'Exports of knowledge-intensive services', 'Sales due to innovation activities', and 'License and patent revenues from selling technologies abroad' (Hollanders & Es-Sadki 2014). Therefore, this paper places a particular focus on indicators to understand the use of evidences in innovation policy.

The influence of indicators in decision-making is largely unknown. Most literature aims to develop indicators, analyse them or evaluate them. However, only a few authors provided clues regarding the extent to which they are used to make a decision.⁴ The existing studies specific to policy contexts revealed that most indicators were often ignored or that their use was limited in policy decisions (MacRae 1985; Lehtonen 2013; Sébastien & Bauler 2013). In 1985, Duncan MacRae argued that the most frequent problem of indicators was their non-use in policy-making. The reasons for this disregard of indicators could be found in the lack of interest, information overload, lack of communication or even opposition to what is being measured (MacRae 1985). More recently, Sébastien and Bauler (2013) emphasized that policy indicators remain largely enigmatic regarding patterns of embeddedness in institutional decision-making processes. In sum, literature about the extent of the use of indicators is meagre and point to a limited use in policy-making.

The literature in innovation contexts has received recent contributions. In fact, Boavida (2015) found that the use of indicators in technology innovation is significantly high (84%), although slightly differentiated in each innovation group: the vast majority of policy makers use indicators (92%), followed close by business R&D&I leaders (89%), and after by (public) researchers (71%). However, social relations were more important than indicators to these decisions for the majority of all decision makers (59%). These results were emphasized by policy makers (68%) and business R&D&I leaders (59%), although half of the researchers (50%) considered indicators as important as social influences. These gaps between the use and the influence of indicators depict the real influences indicators have in the decisions: researchers are more influenced by indicators than business R&D&I leaders and, to a significant extent, than policy makers. These findings confirmed the idea that the use of indicators is different from their influence, as suggested by Gudmundsson and Sørensen (2012) to policy decisions. Therefore, there is a high use of indicators in technology innovation decisions different from their real

³ To the authors, policy-making is the management of rival value set and notions of evidence.

⁴ A significant part of the existing literature about the influence of indicators in policy-making is recent and resulted from two European projects: POINT - Policy Influence of Indicators and PASTILLE - Promoting Action for Sustainability through Indicators at the Local Level in Europe (Bell & Morse 2013).

influence. Furthermore, Boavida (2015) also found that indicators do not play a very significant role in technology innovation: First, indicators had mostly a symbolic role among policy makers (63%) and a limited instrumental role (29%). These results are in line with findings of Gudmundsson and Sørensen (2012), where policy indicators had a very limited direct instrumental role in two sustainable transport policies. However, Boavida (2015) findings disagree with the widespread non-use of indicators in general policy-making, mentioned by MacRae (1985). In fact, only a significant minority of policy makers (8%) revealed that indicators had no role in their decisions. Second, business R&D&I leaders presented a similar but less emphatic pattern. For them indicators had most of the times a symbolic role in the decisions (53%) and a limited instrumental role (36%). Third, the role of indicators to researchers was different. In fact, researchers revealed that indicators can play, almost heterogeneously, an instrumental (35%) and symbolic (35%) role as well as no role (29%) in their decisions. Therefore, indicators play mostly symbolic roles in decisions of policy makers and business R&D&I leaders, but their role with researchers can be more differentiated. Last but not the least, in a study about the significance of composite indicators⁵ for sustainable policy at the EU institutions, composites were found to be not systematically used directly but having an indirect influence on policy-making that needs to be better understood (Sébastien and Bauler 2013). The conclusion emphasized the need to study the processes of evidence-construction, rather than the technical quality of indicators and their independency from their producers (two factors initially presumed important). This latter work, however, dealt only with policy use of composite indicators at EU institutions.⁶ Therefore, there is the need to develop the understanding of the process of construction of evidences more generally to any type of indicators.

The process of evidence construction can help explain the role of indicators play in the decisions. There are two main reasons for this: First, the selection or the disregard of an indicator can be controversial, particularly in contested policy arenas. In fact, “strategic and political use of indicators, manipulation or even abuse of indicators is not necessarily a problem, but rather an essential part of the production of valid and reliable evidence” (Sébastien and Bauler 2013, 10). For example, a significant increase in the number of patents in a country per year can be introduced as an evidence of governmental efforts to promote innovation patents. The example contains the controversial evidence⁷ that governments can directly claim to promote innovation patents, disregarding the efforts of companies and research institutions. If this controversy is brought to the debate, the policy process will determine the influence an indicator can have in providing rational-analytical support to an innovation policy. Second, in policy contexts indicators are used to reduce ambiguity (Sébastien et al. 2014), and may be introduced to reduce the number of variables observed, to simplify and facilitate communication, and to build clear and unambiguous visions of the desired future (Sébastien and Bauler 2013). In these processes, indicators are expected to communicate evidence in a form suited for policy actors that simplify the description of complex systems (Sébastien and Bauler 2013). Therefore, the role of indicators is dependent on their availability and capacity to play a role in the debate that forms the process of construction of evidences, supporting or undermining a policy.

3 METHODOLOGY

The aim of this paper is to understand the use of indicators and other evidences by analysing two processes of construction of evidences in innovation policies. The in-depth analysis of these processes can provide qualitative insights about policy makers’ use of information in practice, how information is valued and the prevalence of formal scientific evidence use in policy decision. The case studies were part of a larger research project aimed to understand the use, influence and role of indicators in decisions specifically of technology innovation. The first case selected was a policy decision to build an electric mobility infrastructure across Portugal. The case of electric mobility is a frequent example

⁵ The aggregation of indicators creates a composite indicator or an index.

⁶ The use of composite indicators were found to be sometimes controversial (Grupp & Schubert 2010; Nardo et al. 2008). In fact, according to some authors composite indicators are considered more adequate to policy communication, rather than to make decisions (Grupp & Schubert 2010).

⁷ It can be said that it also contains a simplification of reality, because innovation efforts can be measure using other evidences than patents.

of innovation in the S&T literature, and a preliminary examination revealed the use of evidences in the media. The second case study was related to the creation of an Iberian nanotechnology laboratory. A preliminary scrutiny revealed a small number of decision makers, geographically accessible in the north of Portugal. The selection of both cases also considered operational restrictions, such as the possibility to review documents and access to potential data and records, as well as the ability to contact and interview decision makers. In addition, it should be noticed that policy makers were a difficult group to investigate because they were less prone to answer surveys, needed substantial explanatory introduction to the research project and revealed the need for more secrecy.

In this context, in-depth interviews were conducted to answer the same open questions about the context of the decision and the process of construction of evidences. They were conducted to cope with the sensitive nature of the information requested, avoid any suspicion of misuse of information, and provide confidentiality to sources when that was possible. They enabled the collection and triangulation of information until saturation was felt; and provided space for other questions to arise and to reveal insights. Furthermore, the interviews included the same guiding questions to analyse the context in relation to the political, economic and organizational environments, and the process of construction of evidences in relation to the way indicators and other evidences were used. The first case study on electric mobility included 9 in-depth interviews to decision makers: 1 to researchers, 4 to business R&D&I leaders and 4 to policy makers. These interviews lasted from one hour up to four hours, and were conducted between February 2011 and March 2013. In the end, two complementary interviews were made to scholars with expertise on the case in March 2012 and in April 2013. The second case about the nanotechnology laboratory included 4 in-depth interviews with decision makers: 2 with researchers and 2 with policy makers. These interviews were conducted in March 2014 and lasted from one hour up to three hours. In the end, one complementary interview was made with a scholar in March 2014.

4 THE ELECTRIC MOBILITY CASE

In 2005, the Portuguese government elected with majority found favourable conditions to engage in the promotion of technological change. In fact, the government supported sound policies towards renewable energies, and believed that they could give a technological push to promote development of the country. At the same time, Portugal was increasingly dependent on costly oil imports that called for measures to de-carbonize the transport sector. Thus, in early 2008, the government decided to create a working group on electric mobility, to develop infrastructure for street charging of EVs across the country. The national programme, hereafter named Mobi-E, was officially launched in mid-2009⁸. Its pilot phase ended in June 2011, with the full implementation of 1300 slow charging posts and 50 fast charging stations in streets, public parking lots, service stations, airports, hotels and shopping centres⁹. A payment system was also implemented to connect personal communication devices (e.g. tablets, smart phones, etc.). By enabling the user to select the most appropriate operation, the system allows for an analysis of mobility costs in order to optimize energy consumption¹⁰. In the end, Mobi-E fully built the infrastructure for charging electric vehicles, but the project failed to address the expected consumers. In fact, only a few cars could be observed using the charging stations in 2012. The charging stations of Mobi-E were supported by the government, through a public innovation support fund created as a counterpart for the granting of wind power licenses (Godinho, Mamede, and Simões 2013). The power company EDP also made initial investments to supply electricity and continues to support the maintenance of the system (costs of around 600 000€/year¹¹).

The decision makers of Mobi-E constructed a rationale to support their decision.¹² A central argument used to justify the decision was that “the lack of a recharging infrastructure deters the acquisition of electric vehicles” (Pinto et al. 2010, p.15). However, two interviewees described that the

⁸ Resolução do Conselho de Ministros n.º 20/2009. Diário da República, 1.ª série — N.º 36 — 20 de Fevereiro de 2009.

⁹ “Electric Mobility - Portugal Showcase to the World – Institutional presentation”. 2010. GAMEP - Gabinete de Apoio à Mobilidade Eléctrica em Portugal.

¹⁰ “Mobi-E Electric Mobility - Portugal Showcase to the World”. Mobi-E. November 2010.

¹¹ According to 2012 costs.

¹² “Modelo de Mobilidade Eléctrica Para Portugal – Apresentação a Sua Excelência O Ministro da Economia e Inovação – Sumário Executivo”. Presentation. Roland Berger Strategy Consultants. Lisboa. 14/1/2009.

decision was not so much based on the technological effect of the policy, but rather on its political and social impact. Furthermore, technical evidences were sought after the decision¹³ was made. The central piece of evidence supporting the Mobi-E decision was based on an indicator of market penetration for EVs in 2020. This indicator was based on an optimistic scenario for the fleet of EVs. In fact, according to the forecast of the coordinator of the office for electric mobility, in 2020 Portugal would have 750000 electric vehicles (Gomes 2010). But, according to a study of Paulo Santos in 2009, there will be no more than 600000 electric vehicles in 2020 in a “very” optimistic scenario (Santos 2009, 40). Moreover, according to Luís Gomes (2010), the governmental forecast was optimistic because it represented 80% of the sales in 2020 (considering a sales growth rate of 1%). In fact, his study forecasted an optimistic scenario with a penetration rate of 50%, predicting only 322 027 electric vehicles in 2020 (Gomes 2010). Therefore, the programme was decided based on political and social considerations, and the evidences used to support the decision were based on optimistic scenarios.

To better understand the optimistic nature of the evidences brought to support the decision, it is also necessary to take into consideration other forecasts. For example, an expert from the Portuguese Automotive Business Association (ACAP) reportedly stated that in a ‘very’ optimistic scenario 300000 vehicles were expected to be sold in the year 2020 (Santos 2009). This forecast implied an optimistic increase both supported in the ratio population/sales of cars existent in countries like Belgium and the Netherlands, as well as in the assumption that in 2020 Portugal will reach these countries’ economic and social development (Santos 2009). Furthermore, there were three other evidences available, although even these proved distant from reality. First, Gomes (2010) short-term calculations for an pessimistic scenario for 2011 and 2012 were above reality. The author forecasted 394 electric vehicles in 2011, but only 193 electric vehicles were sold in Portugal in 2011 (Beltramello 2012 based on Frost & Sullivan 2012). Gomes also calculated 999 electric vehicles in 2012, but there were only around 300 vehicles on the Portuguese roads¹⁴. Second, the most pessimistic scenario of the two pessimistic considered in Santos (2009) study, predicted a meagre presence of electric vehicles in 2020 with only 80000 units. The author described this latter scenario as “catastrophic”, given the “significance of public and private investments expected” to create the infrastructures and fiscal benefits to acquire electric vehicles (Santos 2009, 44). Santos also added that this was a very unlikely scenario, “justified by the non-acceptance of this king of technology in the automotive market” (Santos 2009, 44). Third, two other studies provided further evidences in 2010 and 2011. A study contracted by GALP showed that the penetration forecast of the electric vehicle would be significantly slow.¹⁵ Reiner et al. (2010) forecasted also an optimistic technology scenario where BEVs and fuel cell vehicles will have only 5% of market penetration in 2020 in Europe. Therefore, other available evidences existed before the final implementation of the policy that pointed to a moderation in the expectations about the EV market.

Optimist studies forecasting the advent of the EVs were not a Portuguese unique experience. In fact, Midler & Beaume (2010) reported the existence of three scientific studies in the United States predicting the introduction of EVs. The first one in 1973, elaborated by the Wisconsin University, forecasted a penetration rate of 20% of the total sales in 1980 in the USA. In 1979, a Princeton University study forecasted a slower penetration rate (10%) in 2000. Later, in 1994 the World Resources Institute predicted a 25% penetration rate in the US total sales in 2010. Therefore, the literature describes other scientific studies conducted abroad also based on significantly optimistic scenarios.

There are elements to conclude that other evidences played a role in the decision process. In fact, an international consultancy group and a national consultancy company produced evidences to influence policy-making. In fact, the consultancy group was hired by the government to elaborate a technical report on electric mobility¹⁶, and specified technical features for public charging stations. The report also forecasted an optimistic potential market of 180000 EV and Plug-in Hybrid Electric Vehicle for 2020, with 25000 slow charging public posts and 560 fast chargers. This firm also

¹³ “Modelo de Mobilidade Eléctrica Para Portugal, idem.

¹⁴ “Mobi.europe Newsletter”. Mobi.europe. September 2012.

¹⁵ Interview 7, line 211-214 and TIS.PT (2011).

¹⁶ “Modelo de Mobilidade Eléctrica Para Portugal, idem.

calculated that EVs were 11% more competitive than normal ICE to private owners and 12% to companies. Furthermore, another study was solicited by a firm to a national consultancy company and distributed to influence policy outcomes. Some elements of the study benefited the firms' strategy in the short-term, and influenced policy-making in matters of market-share, norms and regulations related to EVs in Portugal. Forecasts were significantly cautions towards the growth of EV market (TIS.PT 2011), creating controversies about the initial governmental claims. Therefore, there are elements to conclude that evidences were introduced in the policy process to influence the final decision in a controversial context.

To conclude, the Mobi-E programme was mainly based on political and social considerations. The evidences used to support the decision were based on optimistic scenarios. Other available evidences existed before the implementation of the decision, and pointed to a moderation in the expectations about the EV market. Some evidences were privately solicited to companies to influence policy-making process. Furthermore, the case study allows five main conclusions in relation to the process of evidence-construction in policy-making. First, the government used their optimistic forecast as an evidence of good policy, focused on an indicator of market penetration of EVs in 2020, and disregarded other independent evidences. Second, time showed that all forecasts were weak evidences to support the policy decision. Third, the subcontracting of a think-tank to support this policy also produced evidences. Fourth, production of technical knowledge by think-tanks can be used to influence policy by policy makers and companies. Fifth, this case revealed the disputed nature of evidences in policy-making, as discussed in the literature. In fact, what constitutes evidence is debatable, loses strength with controversies, can be brought to debate or ignored if useful, and can be influenced by various parties.

5 THE NANOTECHNOLOGY LABORATORY CASE

The idea to create an international Iberian nanotechnology laboratory, hereafter named INL, was initially defined in a governmental policy briefing, during the preparation of the 2005 Summit between Spain and Portugal. The briefing consisted of half a page with political ideas and technical benefits of the proposal. Both the scientific area and the location of the facilities were intentionally left open.¹⁷ In fact, these definitions would be the result of not only negotiations between the two governments, but also the outcome of discussions among government members. At the time, there were several proposals in various scientific fields to be discussed between both countries, such as nanotechnology, grid computing, biotechnology, biomedicine, energy and risk management. There were also several proposals for different regions in Spain and Portugal. Inside the Portuguese government, the stronger candidates to headquarter the facilities were the border regions of Northern Alentejo, where the Évora Summit would be held, and Braga district where nanotechnology research was stronger. In the end of the Summit, the heads of state agreed to locate the facility in Braga, and nominated a Spanish to be its Director-General.¹⁸ Furthermore, the concept of an Iberian joint research laboratory was well received in both Spanish and Portuguese governmental circles for four main reasons: First, the laboratory would cement relations between countries separated by historical events and not prone to cooperate beyond necessary issues. Second, the cooperation would lead to the creation of the first international research institution in Spain or Portugal.¹⁹ Third, the research facility would be dedicated to an advanced scientific area and an emergent technology. Fourth and last, the facility would be opened to participation of other countries, fostering international collaboration.

¹⁷ Cimeira Luso-Espanhola. 2005. "XXIª Cimeira Luso-Espanhola. Évora, 18 e 19 de Novembro de 2005. Conclusões". Évora. Last accessed in 12/12/2014:

http://www.erse.pt/pt/mibel/construcaoesenvolvimento/Documents/CONCLUS%C3%95ES%20DA%20XXI%20CIIMEIRA%20LUSO-ESPANHOLA%20DE%20EVORA_2005.pdf

¹⁸ Cimeira Luso-Espanhola. 2006. "Conclusões da XXIIª Cimeira Luso-Espanhola Badajoz, 24 e 25 de Novembro de 2006". Badajoz. Last accessed in 12/12/2014:

http://www.erse.pt/pt/mibel/construcaoesenvolvimento/Documents/CONCLUS%C3%95ES%20CIMEIRA_BADAJAZ_2006.pdf

¹⁹ The Institute for Prospective Technological Studies located in Seville was only a European research facility of the Joint Research Centre of the European Commission.

Some evidences were collected during the decision process of the INL. In fact, some elements were found about the international context of investments in nanotechnology, particularly in the USA but also at the EU level (Roco et al. 2000; Roco & Bainbridge 2003; Morrison 2005; Hullmann 2006). Furthermore, Spain conducted a significantly detailed study to determine the activities and necessities in the field and to map and improve technical skills and infrastructures in the period 2005-2010 (Correia et al. 2004). In fact, the study extensively included quantified indicators at regional, national and European level. These indicators included cost of research projects, equipment and their skills; number of researchers and technicians and skills; lists of equipment and projects existent in each laboratory; skills required to operate equipments that already exist, ordered and might be ordered in future; etc. Spain also produced other public reports framing the investments in nanotechnology within the S&T system (Comisión Interministerial de Ciencia y Tecnología 2005a; Comisión Interministerial de Ciencia y Tecnología 2005b). At the time, investments were planned for six Spanish laboratories. To the central government, the INL was part of a larger set of investments that needed to be negotiated with the Spanish regions and their research communities (and later with Portugal). The negotiations required evidences that could be introduced in the assessment of the situation and the distribution of the investments. In addition, Portugal did not produce extensive studies on nanotechnology, despite investments in two new associate laboratories. In fact, only the technical committee preparing the INL creation collected elements to map existing research activities in the country (INL Technical Committee 2006). An interviewee argued that the needs to justify the distribution of investments were lower than in Spain, and mostly directed to the Portuguese nanotechnology community. Moreover, no study was found in both countries that demonstrated an explicit opportunity of investing in nanotechnology and nanoscience versus other scientific areas. In fact, the justifications detected were based on the idea that the USA and other developed countries were investing in this research area.²⁰ However, the same argument is also true for other research areas. Therefore, although both countries introduced evidences in the decision process, the collection of evidences was different in the two Iberian countries: In Spain there were detailed preparatory studies with quantified indicators, and in Portugal there was an ad hoc mapping of research groups.

6 CONCLUSIONS AND DISCUSSION

The results suggest that the process of construction of evidences can help to explain how indicators and other evidences were involved in the decision process. The findings revealed different uses of evidences during the process of construction of evidences. In fact, most evidences were collected to support arguments about the need to implement the policies. In the Mobi-E case, the need for the programme was supported with an indicator of penetration rates of EVs in 2020. Other evidences were also solicited to think-tanks and controversies occurred during the decision process. In the INL case, the evidences were collected with different depth by each country: Portugal mapped existing research activities in the area ad hoc; and Spain had an extensive collection of indicators to negotiate the distribution of investments with regions and the nanotechnology community. Both countries lacked comprehensive evidences to justify the concentration of investment in the field of nanotechnology and nanoscience. Thus, most evidences were collected to provide a rationale to existing policy decisions, although there was an exception in Spain where indicators pre-existed the decision to create the INL. These different uses of evidences are in line with the literature: Flitcroft et al. (2011) signalled an abundance of possibilities for evidence use: in one extreme, evidences can be strictly identified with scientific outputs; in the other end of the spectrum, evidences can be the subjective selection of the available information introduced in an argument to persuade about the truthiness or falsity of a statement. The collection of indicators in Spain reveals a use closer to scientific outputs, whereas the use of the indicator of market penetration in 2020 reveals a use closer to persuasion. Second, the findings are also in accordance to the literature where the strength and quality of an evidence can also be related to the number of controversies that it goes through during its

²⁰ According to Bijker (2014) the history of the Dutch engagement with the identification of nanotechnologies as an important issue for consideration in society, politics and policy makers was done by the Rathenau Institute. A report from this technology assessment institute resulted in getting nanotechnologies on the public agenda, though without any explicit positive or negative undertone. (Bijker 2014)

lifetime (Sébastien and Bauler 2013). In fact, in the Mobi-E case, the indicator of penetration lost much of its strength with the controversies that it went through since its creation. Therefore, these case studies confirmed the literature where a broad type of evidences can be used, and confirm the decrease of strength of indicators due to controversies.

Results also revealed that the role of indicators and other evidences did not particularly increased when business engineers with bachelor degrees and academic scientists with PhDs turned into policy makers. In fact, these policy makers were not particularly engaged in deeper quests for indicators or other evidences then they needed to support their decisions. These findings appear to contradict the Musso and Francioni (2012) idea that the educational level is significantly relevant to the decision-maker response. Alternatively, the results appear to be in line with the literature that described situations as an important factor influencing the role of indicators and evidences: Perri 6 (2002) argued that the situations in which policy makers find themselves shape the information that is selected from the complex set available, and which evidence is rejected or at least downplayed. However, the exception to this was the Spanish collection of indicators found in the INL case, where indicators played a more instrumental role to discuss investments. This suggests that the legitimacy of policy arguments in an adversarial policy context (i.e. regional discussions with the government for investments) depends on the ability of actors to present persuasive analytical evidence, as Sébastien, Bauler, and Lehtonen (2014) recently proposed. In adversarial circumstances, policy makers are more likely to use harder analytical indicators, closer to the concept of scientific evidence, than in a more consensual policy decision.

In sum, the study of the process of construction of evidences helped to explain the way indicators and other evidences are involved in innovation policy. It was possible to conclude that evidences and indicators were brought to decision processes according to their availability and capacity to support the different interests of the actors and the stakeholders. Their role was influenced by the particular situation of the decision makers. More importantly, the use of persuasive analytical evidences appears to be related with the adversity of the policy context. Last, it should be notice that the processes of construction of evidence in policy decisions were significantly different from the scientific process. This is particularly relevant to those that need to deal with both processes. In addition, this angle analysis showed that indicators are one tool among others to support innovation decisions. In fact, the relatively minor instrumental role of indicators suggests that indicators are mostly a complementary instrument of decision. When used relevantly, indicators can offer support to a decision. But there are other significant influences that need to be taken into account to understand the role indicators and other evidences play, such as the social relations of the decision makers and their emotional-intuitive decisions.

ACKNOWLEDGEMENTS

The contribution of Nuno Boavida was financed by Direção Geral de Estatísticas da Educação e Ciência and with a PhD scholarship from the Portuguese Fundação para a Ciência e Tecnologia (Ref. SFRH/BD/76200/2011).

7 BIBLIOGRAPHY

- Bell, S. & Morse, S., 2013, Towards an understanding of how policy making groups use indicators. *Ecological Indicators*, 35, 13–23. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S1470160X12004426> [Accessed October 7, 2014].
- Beltramello, A., 2012, *Market Development for Green Cars*, Paris: OECD Publishing. Available at: http://econpapers.repec.org/paper/oecenvddd/2012_2f3-en.htm [Accessed May 24, 2013].
- Bijker, W., 2014, Technology Assessment: The State of Play - Towards a Hybrid and Pluriform Process of Governance of Science and Technology. In T. Michalek et al., eds. *Technology Assessment and Policy Areas of Great Transitions*. Prague: Informatorium, 23–36. Available at: <http://pacita.strast.cz/en/outputs/outputs-list/proceedings-from-the-pacita-2013-conference-in-prague>.
- Boavida, N., 2015, *Decisions of technology innovation: The role of indicators*. PhD Thesis, Universidade Nova de Lisboa. (forthcoming).

- Boavida, N., Baumann, M., et al., 2013, Technology transition towards electric mobility – Technology Assessment as a tool for policy design. In *GERPISA Conference*. Paris: GERPISA, 1–33. Available at: <http://gerpisa.org/en/node/2085>.
- Boavida, N., Moniz, A.B. & Laranja, M., 2013, Towards an assessment of the Portuguese e-mobility case: The Mobi-E. In T. Michalek et al., eds. *Technology Assessment and Policy Areas of Great Transitions*. Prague: Informatorium, 263–269.
- Comisión Interministerial de Ciencia y Tecnología, 2005a, *Memoria de actividades I+D+I. 2005*, Madrid: Ministerio de Educación y Ciencia. Available at: <http://www.idi.mineco.gob.es/stfls/MICINN/Investigacion/FICHEROS/2005-memoria-i-d-i.pdf>.
- Comisión Interministerial de Ciencia y Tecnología, 2005b, *Plan nacional de investigación científica, desarrollo e innovación tecnológica (2004-2007) - Programa de trabajo para el año 2005*, Madrid: Ministerio de Educación y Ciencia.
- Correia, A., Hernández, J.L.R. & Domingo, P.A.S., 2004, *Estudios de las actividades y necesidades, en el área de las Nanociencias/Nanotecnologías*, Madrid. Available at: http://issuu.com/phantoms_foundation/docs/informe_nanospain_infraestructuras/1.
- Denzin, N.K., 2009, The elephant in the living room: or extending the conversation about the politics of evidence. *Qualitative Research*, 9, 2, 139–160. Available at: <http://qrj.sagepub.com/cgi/doi/10.1177/1468794108098034> [Accessed March 24, 2014].
- Flitcroft, K. et al., 2011, Getting evidence into policy: The need for deliberative strategies? *Social science & medicine (1982)*, 72, 7, 1039–46. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21419539> [Accessed September 19, 2013].
- Freeman, C., 1995, The “National System of Innovation” in historical perspective. *Cambridge Journal of Economics*, 19, 1, 5–24. Available at: <http://cje.oxfordjournals.org/content/19/1/5.short> [Accessed November 14, 2013].
- Gault, F., 2013, *Handbook of Innovation Indicators and Measurement* F. Gault, ed., Cheltenham, UK: Edward Elgar.
- Glasner-Edwards, S. & Rawson, R., 2010, Evidence-based practices in addiction treatment: review and recommendations for public policy. *Health policy*, 97, 2-3, 93–104. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2951979&tool=pmcentrez&rendertype=abstract> [Accessed March 28, 2014].
- Godin, B., 2009, The Culture of Numbers: The Origins and Development of Statistics on Science. Working paper, 40, 1–30.
- Gomes, L.M.M., 2010, *O Veículo Eléctrico e a sua Integração no Sistema Eléctrico*. Master Thesis. Universidade Técnica de Lisboa.
- Grupp, H. & Mogege, M.E., 2004, Indicators for national science and technology - Their Development, Use, and Possible Misuse. In H. F. Moed & E. Al, eds. *Handbook of Quantitative Science and Technology Research*. Kluwer Academic Publishers. Printed in the Netherlands, 75–94.
- Grupp, H. & Schubert, T., 2010, Review and new evidence on composite innovation indicators for evaluating national performance. *Research Policy*, 39, 1, 67–78. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0048733309001991> [Accessed June 10, 2011].
- Gudmundsson, H. & Sørensen, C.H., 2012, Some use—Little influence? On the roles of indicators in European sustainable transport policy. *Ecological Indicators*, 35. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S1470160X12003068> [Accessed September 25, 2012].
- Hall, J.L. & Jennings, E.T., 2010, Assessing the use and weight of information and evidence in U.S. state policy decisions. *Policy and Society*, 29, 2, 137–147. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S1449403510000068> [Accessed October 9, 2013].
- Head, B.W., 2010, Reconsidering evidence-based policy: Key issues and challenges. *Policy and Society*, 29, 2, 77–94. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S1449403510000020> [Accessed September 19, 2013].
- Henn, M., Weinstein, M. & Foard, N., 2009, *A critical introduction to Social Research* 2nd ed., London: SAGE Publications, Inc.

- Hollanders, H. & Es-Sadki, N., 2014, *Innovation Union Scoreboard*, Brussels: European Commission. Available at: http://ec.europa.eu/enterprise/policies/innovation/files/ius/ius-2014_en.pdf [Accessed November 25, 2014].
- Hullmann, A., 2006, *The economic development of nanotechnology - An indicators based analysis*, Brussels. Available at: [http://nanotech.law.asu.edu/Documents/2009/09/The economic development of nanotechnology - an indicators based analysis_220_7306.pdf](http://nanotech.law.asu.edu/Documents/2009/09/The_economic_development_of_nanotechnology_-_an_indicators_based_analysis_220_7306.pdf) [Accessed April 16, 2014].
- INL, 2011a, Creating value at nanoscale. 1–12.
- INL, 2011b, *INL – Laboratório Ibérico Internacional de Nanotecnologia - Novembro de 2009 a Maio de 2011 - Atualização da informação sobre a criação e actividades do INL publicada no volume sobre as actividades do Ministério da Ciência, Tecnologia e Ensino Superior, 2*, Braga.
- INL, 2012, International Iberian Nanotechnology Laboratory - An international community for nanotechnology.
- INL Technical Committee, 2006, *Portugal-Spain International Research Laboratory - International Iberian Nanotechnology Laboratory (INL)*, Braga. Available at: [http://www.unic.pt/images/stories/publicacoes4/80 Technical Committee Report.pdf](http://www.unic.pt/images/stories/publicacoes4/80_Technical_Committee_Report.pdf).
- Jauch, L. & Glueck, W., 1988, An invitation to strategic management. In L. Jauch & W. Glueck, eds. *Business policy and strategic management*. Singapore: McGraw-Hill, Inc, 1–42.
- Jonson-Reid, M., 2011, Disentangling system contact and services: A key pathway to evidence-based children's policy. *Children and Youth Services Review*, 33, 5, 598–604. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0190740910003440> [Accessed March 28, 2014].
- Juntti, M., Russel, D. & Turnpenny, J., 2009, Evidence, politics and power in public policy for the environment. *Environmental Science & Policy*, 12, 3, 207–215. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S1462901108001391> [Accessed September 19, 2013].
- Lehtonen, M., 2013, The non-use and influence of UK Energy Sector Indicators. *Ecological Indicators*, 35, 24–34. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S1470160X12003779> [Accessed October 22, 2014].
- Lilienfeld, S.O. et al., 2013, Why many clinical psychologists are resistant to evidence-based practice: root causes and constructive remedies. *Clinical psychology review*, 33, 7, 883–900. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23647856> [Accessed March 25, 2014].
- Lomas, J. et al., 2005, *Conceptualizing and Combining Evidence for Health System Guidance - Final Report*, Ottawa. Available at: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.112.4445> [Accessed December 18, 2013].
- Luhmann, N., 1990, Technology, environment and social risk: a systems perspective. *Industrial Crisis Quarterly*, 4, 3, 223–231. Available at: <http://oae.sagepub.com/cgi/doi/10.1177/108602669000400305> [Accessed May 28, 2013].
- MacRae, D., 1985, *Policy indicators : links between social science and public debate*, Chapel Hill. Available at: <http://searchworks.stanford.edu/view/1129898>.
- Mamede, R.P., Simões, V.C. & Godinho, M.M., 2014, Assessment and challenges of industrial policies in Portugal: is there a way out of the “stuck in the middle” trap? In Aurora Teixeira, E. Silva, & R. P. Mamede, eds. *Structural Change, Competitiveness and Industrial Policy Painful Lessons from the European Periphery*. London: Routledge, 1–21.
- Midler, C. & Beaume, R., 2010, Project-based learning patterns for dominant design renewal: The case of Electric Vehicle. *International Journal of Project Management*, 28, 2, 142–150. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0263786309001173> [Accessed April 29, 2013].
- Morrison, M., 2005, *European Nanotechnology Infrastructure and Networks*, Brussels. Available at: <http://www.nanowerk.com/nanotechnology/reports/reportpdf/report54.pdf>.
- Musso, F. & Francioni, B., 2012, The Influence of Decision-Maker Characteristics on the International Strategic Decision-Making Process: An SME Perspective. *Procedia - Social and Behavioral Sciences*, 58, 279–288. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S1877042812044643> [Accessed September 19, 2013].

- Mytelka, L. & Smith, K., 2002, Policy learning and innovation theory: an interactive and co-evolving process. *Research policy*, 31, 8-9, 1467–1479. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0048733302000768> [Accessed April 8, 2014].
- Nardo, M. et al., 2008, *Handbook on Constructing Composite Indicators: Methodology and User Guide* OECD & J. E. Commission, eds., Paris: OECD Publishing. Available at: <http://www.oecd-ilibrary.org/oecd/content/book/9789264043466-en>.
- Pandza, K., Wilkins, T. & Alfoldi, E., 2011, Collaborative diversity in a nanotechnology innovation system: Evidence from the EU Framework Programme. *Technovation*, 31, 9, 476–489. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0166497211000721> [Accessed April 9, 2014].
- Perri 6, 2002, Can Policy Making be Evidence-Based? *MCC: Building knowledge for integrated care*, 10, 1, 3–8.
- Pinto, M., Reis, L. & Neves, T., 2010, MOBi.E.-The Portuguese Programme for Electric Mobility. In *European Transport Conference, 2010*. Glasgow, United Kingdom, 1–19. Available at: <http://abstracts.aetransport.org/paper/index/id/3437/confid/16> [Accessed April 25, 2013].
- Porter, C., 2010, *What Shapes the Influence Evidence Has on Policy?*, Working Paper, 62, Young Lives, Department of International Development, University of Oxford, Oxford.
- Reiner, R. et al., 2010, *Challenges for a European market for electric vehicles*, Brussels. Available at: <http://www.europarl.europa.eu/document/activities/cont/201106/20110629ATT22885/20110629ATT22885EN.pdf> [Accessed May 23, 2013].
- Roco, M.C. et al., 2000, *National nanotechnology initiative: The initiative and its implementation plan* M. C. Roco & W. S. Bainbridge, eds., Washington, D.C.: National Science Foundation. Available at: http://nano.gov/sites/default/files/pub_resource/nni_implementation_plan_2000.pdf [Accessed June 4, 2014].
- Roco, M.C. & Bainbridge, W.S.W.S., 2003, *Nanotechnology: Societal Implication - Maximizing Benefit for Humanity*, Arlington, VA. Available at: http://www.nsf.gov/crssprgm/nano/reports/nni05_si_societal_implications_2005.pdf.
- Santos, P.M.C., 2009. *O Automóvel Híbrido como Elemento Fornecedor / Consumidor de Electricidade - Modelação de Baterias de Iões de Lítio*. Master Thesis. Universidade Técnica de Lisboa.
- Sébastien, L. & Bauler, T., 2013, Use and influence of composite indicators for sustainable development at the EU-level. *Ecological Indicators*, 35, 3–12. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S1470160X13001775> [Accessed October 3, 2014].
- Sébastien, L., Bauler, T. & Lehtonen, M., 2014, Can Indicators Bridge the Gap between Science and Policy? An Exploration into the (Non)Use and (Non)Influence of Indicators in EU and UK Policy Making. *Nature and Culture*, 9, 3, 316–343. Available at: <http://openurl.ingenta.com/content/xref?genre=article&issn=1558-6073&volume=9&issue=3&spage=316> [Accessed October 22, 2014].
- Sorrell, S., 2007, Improving the evidence base for energy policy: The role of systematic reviews. *Energy Policy*, 35, 3, 1858–1871. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0301421506002667> [Accessed October 9, 2013].
- Stacey, R.D., 2011, *Strategic management and organisational dynamics - The challenge of complexity* 6th ed., Harlow: Pearson Education Limited.
- TIS.PT, 2011, *Transportes, energia e fiscalidade: Desafios e oportunidades perante a evidência da mudança*, Lisboa, Portugal.
- Veltri, G., Lim, J. & Miller, R., 2014, More than meets the eye: the contribution of qualitative research to evidence-based policy-making. *Journal of Social Science Research*, 27, 1, 1–4. Available at: <http://www.tandfonline.com/doi/abs/10.1080/13511610.2013.806211> [Accessed March 28, 2014].
- Weiss, C.H., 1979, The Many Meanings of Research Utilisations. *Public Administration Review*, September/October, 426–431.



Evidence-based policy learning: the case of the research excellence indicator

Sjoerd Hardeman* and Daniel Vertesy**

*s.hardeman@cpb.nl

CPB The Netherlands Bureau for Economic Policy Analysis, The Hague, The Netherlands

** daniel.vertesy@jrc.ec.europa.eu

Unit of Econometrics & Applied Statistics, Joint Research Centre, European Commission, Ispra, Italy

Abstract

Excellence is arguably the single most important concept in academia today, especially when it comes to science policy making. At the same time, however, excellence leads to a great amount of discomfort, leading some to plea for an overall rejection of the concept. The discomfort with excellence reaches its heights whenever proposals are made for measuring it. Yet, especially given the period of professionalization science policymaking finds itself in, these same metrics are frequently called upon to legitimate policy interventions. Excellence and its measurement, it seems therefore, is something we can neither live with nor without. This paper offers some middle ground in the debate on excellence and its measurement for science policy purposes. Using the case of the European Commission's Research Excellence Indicator as an example, we show how the development and use of indicators offers an opportunity for learning in science policy making. Ultimately, therefore, we show how and in what ways measuring excellence can contribute to evidence-based science policy learning in practice.

Introduction

In an evidence-based society (Smith, 1996), indicators are increasingly called upon to inform policy (Porter, 2015). Indicators are seen as viable tools to equip policymakers with the information required to arrive at informed decisions. On the one hand, then, indicators are appealing as they simplify complex phenomena into comprehensible and authoritative information (Porter, 1996; Espeland, 2015). On the other hand, however, the popularity of indicators for policy purposes has evoked considerable criticisms as well (cf. Kelley and Simmons, 2014). A number of studies started to assess the implications of indicators on ways of knowing and governing society (Davis et al., 2012; Fioramonti, 2014; Rottenburg et al., 2015). However, few attempts have been made thus far to understand how such indicators come about in practice. Nevertheless, such attempts are crucial for understanding how the work on indicators done "back stage" feeds into what becomes visible "front stage" in the policy domain (Nowotny, 2007, p. 484). This paper seeks to uncover and reflect upon the work usually left invisible.

We use the European Commission's research excellence indicator (Hardeman et al., 2013) to illustrate how indicators-for-policy come about in practice. On the one hand, research excellence is perhaps the most celebrated topic amongst research policymakers today. Not least amongst policymakers operating at the supranational level. For example the OECD recently came with a report on "promoting research excellence" (OECD, 2014) and the World Bank, likewise, discusses "the road to academic excellence" (Altbach & Salmi, 2011). On the

other hand, the notion of research excellence is heavily contested amongst research's main constituents; that is, researchers. Some have even called for an altogether rejecting of the notion of research excellence (Stilgoe, 2015). In this paper we seek to show how, in the midst of both such appraisal and controversy, quantifying research excellence has come about within the EU policy context. As we have been actively involved in the construction of this research excellence indicator, we believe we are in a unique position to offer such a reflexive account.

Indicators-for-policy: principles and challenges

Indicators are usually thought of as numeric representations of the real world around us. More formally, Davis et al. (2015) define the indicator as “a named collection of rank-ordered data that purports to represent the past or projected performance of different units.” The surge of indicators-for-policy can be explained by internationalization and globalization trends on the one hand, accompanied by calls for accountability on the other. First, internationalization and globalization trends are likely to have increased policymakers' demand for information that is comparable across distant settings. Indicators meet this demand in that they render distant objects (countries, regions, organizations, people) comparable. Second, alongside this internationalization and globalization trend there has been an increasing call for accountability or what some have called an “audit explosion” (Power, 1994). This call for accountability has at least two dimensions. One is that decision makers, especially in the public domain, have an incentive to demonstrate their success objectively *ex post*. The other is that decision makers have an incentive to legitimize their activities: calling for information in which to ground their decisions *ex ante*. Overall, indicators promise to facilitate rational – evidence-based – policymaking by offering unbiased, comparable information on a single phenomenon of interest for different units.

Indicators seek to capture a semiotic relationship between the signified on the one hand and the signifier on the other (Boulanger, 2014). It follows that, in principle, the extent to which the signifier (natural system) matches (is encoded into and decoded from) the sign (formal system) determines the validity of the indicator. Substantiating the relationship between the signified natural system and the signifying formal system usually involves an act of quantification. Quantification means “to convert into numerical existence what was previously expressed in words and not in numbers” (Desrosieres, 2015). Quantification is different from measuring in that whilst the former brings into existence something that previously did not exist, the latter implies that something existed all along and therefore can be readily measured. Like any other modeling activity then, quantification is a craft in which the craftsman makes choices based on theoretically informed judgments constrained by practical feasibility.

The problem with indicators as the outcome of quantification activities is that indicators focus on some aspects of the real world around us and leave other aspects out (Espeland, 2015). Indicators offer only a partial view on the real world. Indicators are not a given; rather, they are actively constructed. A major issue of quantification thus holds that the outcome of indicators is suspect of being steered by the particular assumptions adopted in its construction. For example, in the context of Science & technology indicators, Grupp and Mogege (2004, p. 1378) argue that “Considerable room exists for manipulation by selecting, weighting and aggregating indicators.” Indeed, as shown by their empirical analysis, by tweaking the underlying assumptions of S&T indicators, scores and rankings change considerably (Grupp and Mogege, 2004; Grupp and Schubert, 2010).

The role of normative deliberation in the construction of indicators-for-policy puts severe limits on the view of indicators offering (objective) information allowing policymakers to arrive at (rational) policy decisions. Indeed, the normative element of indicator construction easily turns evidence-based policymaking into policy-based evidence making (Strassheim & Kettunen, 2014). Accordingly, Barré (2010, p. 228), specifically talking about science & technology indicators, takes indicators as the outcome of negotiations: “the quantification process is of a political nature”. What is more, in his view, the normativity involved in the construction of indicators-for-policy renders them to operate like debatable devices (Barré, 2010). The notion of indicators as debatable devices, however, does not mean that they are of no use throughout the policy process. On the contrary, for Barré (2001) indicators should be taken “not as a final result to be accepted, but as an entry point for debate.” Beyond indicators as simply debatable devices, they might hence best be taken as debatable debating devices; recognizing both their communicative and instrumental uses.

Quantifying research excellence in practice

The construction of the Research Excellence Indicator

The expert group on the measurement of innovation set up by the European Commission’s Directorate-General Research & Innovation (DG-RTD) was requested “to reflect on the indicators which are the most relevant to describe the progress to excellence of European research” (Barré et al., 2011, p. 3). At that point the whole notion of excellence was said to be “in a rather fuzzy state” (Barré et al., 2011, p. 3). In order to overcome the conceptual confusion surrounding research excellence and to come up with a short list of indicators capable of grasping research excellence, the expert group proceeded in four steps. First, they defined and described types of activities eligible for being called excellent. Second, a set of potential indicators were identified. Third, from this set of potential indicators a short list of (actually available) indicators was recommended. And fourth, a process for interpreting research excellence as a whole at the level of countries was proposed. In all, recognizing that the complete set of indicators needed to be reduced as to become interpretable by policymakers, the expert group thus proposed to come up with 6 distinct composite indicators capturing distinct issues of excellence; that is, excellence in research, excellence in innovation, excellence through impact, excellence through openness, and excellence through attractiveness (Barré et al., 2011).

In a next step, Vertesy & Tarantola (2012) explored the possibility to develop a single (rather than six) composite indicator(s) measuring research excellence. Whilst building upon the framework offered by the expert group, they ultimately endorsed a different set-up. The main issue with the framework proposed by the expert group was that indicators underlying each dimension did not add up. In so doing, Vertesy & Tarantola (2012) came up with a proposal for a single composite indicator measuring research excellence that closely resembles the theoretical framework offered by the expert group whilst at the same time being statistically sound (i.e. one that is coherent from a statistical perspective).

Presented at a workshop organized in Ispra (Italy) during fall 2012 by the European Commission and attended by both policymakers and academic scholars, the newly proposed composite indicator met with fierce criticism. A first critique raised was that the proposed composite indicator mixes up both inputs and outputs whilst research excellence, supposedly, should be about research outputs only. The main point holds that, whereas the outcomes of research and innovation activities are fundamentally uncertain, the nature and magnitude of research and innovation inputs say little to nothing about their outputs. A second critique

raised during the workshop was that some of the indicators used, whilst certainly pertaining to research, need not say may much about their excellent content. In as much as the underlying indicators were referring to outputs, their characterization should pertain to the best outputs only, therewith by and large excluding all underlying indicators that referred to any kind of input (e.g. gross investment in R&D) or any kind of process organizing the translation of inputs into outputs (e.g. university-industry collaborations). The main point to emphasize here is that research excellence, according to these two critiques at least, is first and foremost about outputs and, secondly, not just any kind of output but only those outputs meeting the highest quality standards.

Taking these critiques on board, the research excellence indicator has been further refined towards the finalization of the 2013 report (Hardeman et al., 2013). First, the scope of the indicator was made explicit by limiting it to research in science and technology only. Second, following up on the critique strongly suggesting to distinguish inputs from outputs, and conditional upon being available for both codified and tacit knowledge, primarily those underlying indicators were considered that focused on outputs. Third, from those outputs, only those were taken into account that explicitly made reference to high-quality aspects of research in science and technology. Ultimately, then, the theoretical framework for the composite indicator on research excellence was made of a single pillar. Given that the underlying indicators are not available for all countries, the rankings presented in the 2013 Innovation Union Competitiveness Report are based on a single composite indicator aggregating either three (non-ERA countries) or four (ERA countries) underlying indicators (European Commission, 2013).

Current state of play

Soon after the publication of the research excellence indicator in 2013 and the inclusion of its results in the 2013 Innovation Union Competitiveness Report a new request was made by DG-RTD to update and, if deemed necessary, refine the indicator for future reports. This request for an update led to a re-assessment of the existing indicator (Hardeman & Vertesy, 2015). Given the methodological focus of the reassessment, we chose to use uncertainty- and sensitivity analysis to address the implications of constructing the indicator based on a set of different underlying assumptions (Hardeman & Vertesy, 2015).

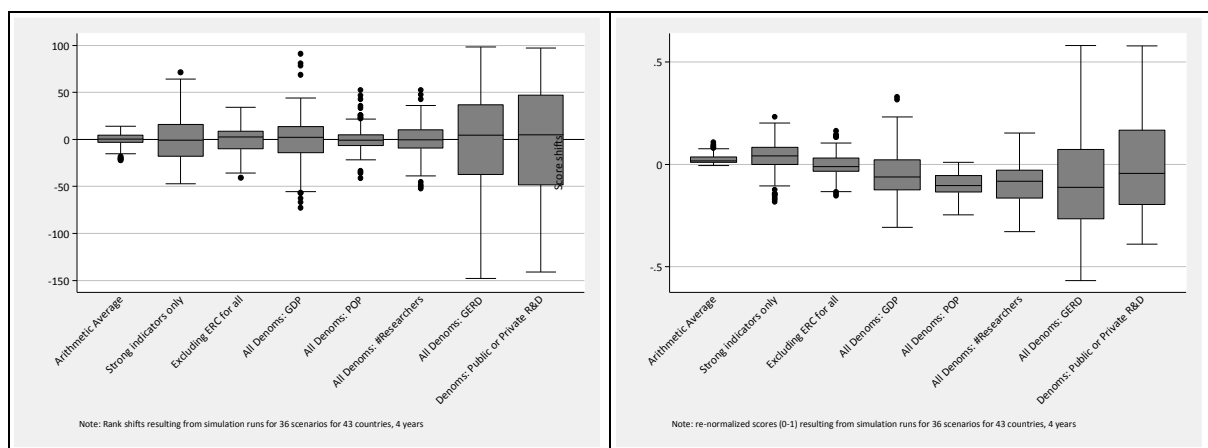
Table 1. Uncertainties and alternative modelling choices addressed in the uncertainty and sensitivity analysis of the Research Excellence Indicator

Reference (baseline scenario)	Alternatives
<i>Uncertainty in the aggregation formula</i>	
geometric average	arithmetic average
<i>Uncertainty in the selection of component indicators</i>	
HICIT, PCT, TOPINST, ERC	Strong indicators only (HICIT, PCT) Exclude ERC for all countries (HICIT, PCT, TOPINST)
<i>Uncertainty in the selection of denominator for rescaling components</i>	
Indicator-specific denominators (HICIT / Total Publication; PCT/Population; TOPINST/Population; ERC/Public R&D)	All components denominated by GDP
	All components denominated by Population
	All components denominated by Nr. Researchers
	All components denominated by GERD
	Components denominated by PCT/BERD, all others divided by Public R&D (GOVERD+HERD)

Source: Hardeman & Vertesy (2015)

Table 1 offers an outline of the various methodological alternatives considered throughout the uncertainty- and sensitivity analysis and **¡Error! No se encuentra el origen de la referencia.** offers the main outcomes of this analysis. Clearly, some methodological choices have a larger effect than others. In particular two results stood out. One is about the choice of numerators: restricting the set of numerators to established variables like the number of highly cited publications and the number of PCT patents greatly effects volatility in scores and ranks. The other is about the choice of denominators: including either monetary denominators (like R&D expenditures) or non-monetary denominators (like population) also greatly influences the scores and ranks one ends up with. The main message put forward in conclusion is that any indicator measuring research excellence (at the country level) “*crucially depends upon the basic conceptual framework underlying it*” (Hardeman & Vertesy, 2015).

Figure 1. Results of the sensitivity analysis on the research excellence indicator: average shifts in country ranks (left panel) and scores (right panel) due to changing modeling assumptions



Source: Hardeman and Vertesy (2015)

In a way, the findings of the uncertainty- and sensitivity analyses highlighted the importance of normative choices (Hardeman & Vertesy, 2015). With regards to the challenges addressed, however, such normative debates materialized only within a limited arena. This happened in the context of a direct request by the European Research Area and Innovation Committee (ERAC) – a strategic policy advisory committee consisting of Member States and the EU Commission – to use a modified version of the Research Excellence indicator to measure country progress towards ERA roadmap priority 1 (“effective national research systems”). The ERAC was rather specific in its request to replace the component measuring top universities and research organizations with an indicator on Marie Curie fellowships, while keeping the other three indicators of the Research Excellence composite unchanged. This new component would add to the composite index the aspect of “capacity building” as well as “how well organized and attractive a research system already is when dealing with people who represent the future of research and innovation in Europe” (Vertesy, 2015). The modified indicator was subsequently adopted by the ERAC and the Competitiveness Council. The choice of the ERAC represents an interesting new interpretation of research excellence in the context of effectiveness. While the ERAC was in favor of keeping the composite indicator with a single pillar, the choice also represents a departure – limited it may be – from the narrow understanding of research excellence.

Summing up a personal history

From our personal history of the research excellence indicator we can identify several turning points that, at least in our view, have left important marks on the way its construction has

unfolded (see table 2). First, the Barré et al. (2011) report clearly set the scene for what was to come. Whilst policymakers, perhaps from a concern with fiscal stringency were primarily concerned with what gets into and comes out of research, Barré et al. (2011) clearly emphasized the processes underlying research as well.

Table 2 Turning points in the construction of the research excellence indicator

Occasion	Context	Result
2011 Barré et al. report ⁷	EC aims to collect indicators in order to monitor the development of the Innovation Union Flagship Initiative of the EU2020 strategy “...The expert group will be requested to reflect on the indicators which are the most relevant to describe the progress to excellence of European Research’ and make ‘recommendation of a choice of short list of indicators justified by economic and policy facts and figures”.	Proposal of a conceptual framework involving knowledge production (impact) and institutional arrangements (openness, attractiveness), seen through the engagement of... <ul style="list-style-type: none"> • Research actors • Industrial innovation actors and • Societal and political actors
2012 Vértésy & Tarantola report	Feasibility study on a composite indicator measuring research excellence based on the Barré et al. (2011) framework	<ul style="list-style-type: none"> • Difficult to operationalize the Barré et al framework; • Alternative frameworks tested with the aim to strengthen statistical coherence & robustness
2012 October Workshop	Validation of outcomes (note: Barré not present)	<ul style="list-style-type: none"> • Message of Workshop attendees: better focus; distinguish inputs, interactions and outputs and focus on the latter
2013 Hardeman et al. (2013) report	<ul style="list-style-type: none"> • Address the points of criticism • Proposal for a revised conceptual framework ‘national research systems’ 	<ul style="list-style-type: none"> • Sharpened definition of research excellence • Proposal of a composite • Responds to policy request on the choice of years
2015 Hardeman & Vértésy report	<ul style="list-style-type: none"> • Sørensen et al (2015) critique • On the occasion of the latest indicators update, address issues of list of indicators (strong versus weak; alternative numerators and denominators), weighting and aggregation and time coverage 	<ul style="list-style-type: none"> • Sensitivity analysis highlighting the importance of denominators; no single best option can be discerned from technical operations only

Second, however, the importance of the Barré et al. (2011) report clearly does not reside in its widespread endorsement. For any policy maker who may have found the number of indicators proposed by the Barré et al. (2011) report discomfoting, the conclusion of Vértésy & Tarantola (2012) that the variables did not add up to point at a single latent dimension may have come as a relief. Note then that the technical analysis offered by Vértésy & Tarantola (2012) was instrumental and therewith paved the way for the further unfolding of the research excellence indicator.

Third, the indicator (or set of indicators) initially proposed by Barré et al. (2011) was further stripped in response to criticisms raised during the October 2012 workshop. Based on only four components, the indicator presented in the 2013 Innovation Union Competitiveness Report (European Commission, 2013) can be considered a direct outcome of the criticisms raised against earlier proposals during the October 2012 workshop.

In conclusion, this did not settle the debate however. As with the coherence analysis offered by Vertesy & Tarantola (2012), technical analyses once more seem to have been instrumental for steering the debate on the research excellence indicator into new directions. However, whilst before such technical analysis served the interest of policymakers to come up a single instead of a battery of indicators, now it seems to have opened the debate on the future of the research excellence indicator, which – at least for now – was settled with its use for the purposes of the ERAC.

Concluding remarks

Focusing on the research excellence indicator on offer from the European Commission, this paper sought to uncover and reflect upon the work typically left invisible in the production of indicators. Although indicators-for-policy are usually endorsed with reference to evidence-based policy making, the history of the research excellence indicator shows that politics are as much an impetus to as it is the outcome of the production of indicators. First, the production of indicators as a socio-political process resonates Barré's (2010) characterization of science & technology indicators as *debatable* devices. It has been illustrated that constructing indicators inevitably runs into making choices that ultimately can only be legitimized with reference to normative considerations. Such normative underpinnings render the indicator open for debate. Also, whilst Barré (2010) offers a normative account on how indicator construction should take place, our case study offers a positive account on how indicator construction has taken place in accordance with the idea of indicators as debatable devices.

Second, however, beyond debatable devices, the production of indicators, including the use of uncertainty- and sensitivity analysis therein, also evokes an ideal of indicators as *debating* devices. Indicators not only have normative underpinnings rendering them debatable, they can also be used to trigger and structure debate. Indeed, the use of such technicalities can be very instructive, offering insights on the scope of quantifying research excellence for (evidence-based) policy purposes. Such an instructive use of technicalities, however, has two faces: technicalities can be used strategically towards both closing and opening the debate on the normative underpinnings of indicators.

In conclusion, we believe that conceiving of indicators as debatable debating devices points at a role for indicators in policy that is not so much about evidence-based policy making rather than evidence-based policy learning. In contrast to evidence-based policy making, emphasis is much more on the process nature of evidence and policy. Neither evidence nor policy should be taken as products. Indicators as debatable debating devices have a role to play here as they offer both policymakers and the (lay) public a tool to hold on to throughout the formulation and reformulation of policies.

References

- Altbach, P. G., & Salmi, J. (Eds.). (2011). *The road to academic excellence: The making of world-class research universities*. World Bank Publications.
- Barré, R. (2001). Sense and nonsense of S&T productivity indicators. *Science and Public Policy*, 28(4), 259-266.
- Barré, R. (2010). Towards socially robust S&T indicators: indicators as debatable devices, enabling collective learning. *Research Evaluation*, 19(3), 227-231.

Barré, R. Hollanders, H. & Salter, A. (2011). Indicators of research excellence. Expert Group on the measurement of innovation. Luxembourg: Publ. Office of the European Union

Boulanger, P.-M. (2014) “Elements for a comprehensive assessment of public indicators”. *JRC Scientific and Policy Reports*, Luxembourg: Publ. Office of the European Union.

Bristow, G. (2005). Everyone's a ‘winner’: problematising the discourse of regional competitiveness. *Journal of Economic Geography*, 5(3), 285-304.

Cornell University, INSEAD, & WIPO (2014). *The Global Innovation Index 2014: The Human Factor In Innovation*. Fontainebleau, Ithaca, and Geneva: WIPO and CII

Davis, K., Fisher, A., Kingsbury, B., and Merry, SE. (2012). *Governance by indicators: global power through classification and rankings*. Oxford: Oxford University Press.

Desrosieres, A. (2015). Retroaction: how indicators feed back onto quantified actors. In: Rottenburg et al. (eds). *The world of indicators: the making of governmental knowledge through quantification*. Cambridge: Cambridge University Press.

Espeland, W. (2015). “Narrating numbers”. In: Rottenburg et al. (eds). *The world of indicators: the making of governmental knowledge through quantification*. Cambridge: Cambridge University Press.

European Commission (2000). *Towards a European Research Area*, COM(2000) 6 Final, Brussels, 18.01.2000.

European Commission (2007). *Green Paper, The European Research Area: New Perspectives*. COM(2007) 161., Brussels, 4.4.2007.

European Commission (2011). *Innovation Union Competitiveness Report. 2011 edition*. Luxembourg: Publications Office of the European Union.

European Commission (2012). *A Reinforced European Research Area Partnership for Excellence and Growth*. COM(2012) 392 final, Brussels, 17.7.2012.

European Commission (2014). *Innovation Union Competitiveness Report 2013*. Luxembourg: Publications Office of the European Union.

European Council (2000). Lisbon European Council 23 and 24 March 2000. Presidency Conclusions.

Fioramonti, L. (2014). *How numbers rule the world: the use and abuse of statistics in global politics*. Zed Books.

Freeman, C. & Soete, L. (2009) Developing science, technology and innovation indicators: What we can learn from the past. *Research Policy*, 38(4) 583–589

Grupp, H., & Mogege, M. E. (2004). Indicators for national science and technology policy: how robust are composite indicators?. *Research Policy*, 33(9), 1373-1384.

Grupp, H., & Schubert, T. (2010). Review and new evidence on composite innovation indicators for evaluating national performance. *Research Policy*, 39(1), 67-78.

Hardeman, S., Van Roy, V., Vertesy, D., & Saisana, M. (2013). *An analysis of national research systems (I): A Composite Indicator for Scientific and Technological Research Excellence*. Luxembourg: Publications Office of the European Union. doi: 10.2788/95887

INSEAD (2013): *The Global Talent Competitiveness Index 2013*. Singapore: INSEAD and HCLI.

Kelley, J. G., & Simmons, B. A. (2015). Politics by number: indicators as social pressure in international relations. *American Journal of Political Science*, 59(1), 55-70.

King, D. A. (2004). The scientific impact of nations. *Nature*, 430(6997), 311-316.

May, R. M. (1997). The scientific wealth of nations. *Science*, 275(5301), 793.

Nowotny, H. (2007). How many policy rooms are there? Evidence-based and other kinds of science policies. *Science, Technology & Human Values*, 32(4), 479-490.

OECD (2000). *Science, Technology and Industry Outlook 2000*, Paris: OECD Publishing.

OECD (2013). *OECD Science, Technology and Industry Scoreboard 2013: Innovation for Growth*, Paris: OECD Publishing.

OECD (2014), *Promoting Research Excellence: New Approaches to Funding*, OECD Publishing. <http://dx.doi.org/10.1787/9789264207462-en>

Pielke, R. A. (2007). *The honest broker: making sense of science in policy and politics*. Cambridge: Cambridge University Press.

Porter, T. M. (1996). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton University Press.

Porter, T.M. (2015). “The flight of the indicator”. In: Rottenburg et al. (eds). *The world of indicators: the making of governmental knowledge through quantification*. Cambridge: Cambridge University Press.

Power, M. (1994). *The audit explosion* (No. 7). Demos.

Rottenburg, R., Merry, S. E., Park, S. J., & Mugler, J. (eds.). (2015). *The World of Indicators: The Making of Governmental Knowledge Through Quantification*. Cambridge University Press.

Rottenburg, R., Merry, S. E. (2015). “Introduction: a world of indicators: the making of governmental knowledge through quantification”. In: Rottenburg et al. (eds). *The world of indicators: the making of governmental knowledge through quantification*. Cambridge: Cambridge University Press.

Saltelli, A. (2007). Composite indicators between analysis and advocacy. *Social Indicators Research*, 81(1), 65-77.

Saltelli, A., D’Hombres, B., Jesinghaus, J., Manca, A. R., Mascherini, M., Nardo, M., & Saisana, M. (2011). “Indicators for European Union Policies. Business as usual?” *Social Indicators Research*, 102(2), 197-207. 15

Slaughter, S., & Rhoades, G. (1996). The emergence of a competitiveness research and development policy coalition and the commercialization of academic science and technology. *Science, Technology & Human Values*, 21(3), 303-339.

Smith, A. F. (1996). Mad cows and ecstasy: chance and choice in an evidence-based society. *Journal-Royal Statistical Society Series A*, 159, 367-384.

Stilgoe, J. (2014). Against excellence. *The Guardian*, December, 19

Strassheim, H., & Kettunen, P. (2014). When does evidence-based policy turn into policy-based evidence? Configurations, contexts and mechanisms. *Evidence & Policy: A Journal of Research, Debate and Practice*, 10(2), 259-277.

Sørensen, M. P., Bloch, C., & Young, M. (2015). Excellence in the knowledge-based economy: from scientific to research excellence. *European Journal of Higher Education*, 1-20. doi:10.1080/21568235.2015.1015106

Tijssen, R. J. (2003). Scoreboards of research excellence. *Research Evaluation*, 12(2), 91-103.

UNDP (2014). *Human Development Report 2014 – Sustaining Human Progress: Reducing Vulnerabilities and Building Resilience*. New York: United Nations Development Programme.

Vertesy, D. (2015). *Technical note on the robustness of a modified research excellence index*. Luxembourg: Publ. Office of the European Union. doi:10.2760/81642

Vertesy, D., & Tarantola, S. (2012). *Composite Indicators of Research Excellence*. Luxembourg: Publ. Office of the European Union. doi:10.2788/45492



Who sets up the bridge? Tracking scientific collaborations between China and the EU28

Lili Wang* and Xianwen Wang**

*wang@merit.unu.edu

UNU-MERIT, Boschstraat 24, Maastricht, 6211 AX (The Netherlands)

** xianwenwang@dlut.edu.cn

WISE Lab, Dalian University of Technology, Dalian 116085 (P.R.China)

ABSTRACT

In the past decade, collaborations between China and European Union have been rapidly expanding. Hitherto, however, little research has been carried out to assess implementation and impacts of such collaborations. This paper evaluates the collaboration performance and impacts of such collaborations. This paper evaluates the collaboration performance between China and the EU28 concerning major research and innovation priorities. To shed light on the initiatives of collaborations, corresponding authors are detected and classified into three categories, Chinese local, Chinese abroad, and non-Chinese. In order to foster more profitable collaboration for both parties and to formulate options for international policy on research and innovation cooperation between the European Union and China, this paper presents an in-depth analysis of the scientific collaborations focusing on the initiatives and benefits of the collaborations.

1. INTRODUCTION

Along with its fast growth, China has acknowledged the importance of international collaboration for the fulfilment of research and innovation objectives. A number of bilateral and multilateral cooperation agreements and programmes with different countries have been established to stimulate knowledge transfer across national borders. This has served to strengthen formal collaborations and to enhance the scope for deepened institutional cooperation. The 12th Five Year Plan for Science and Technology Development stated that the internationalisation of scientific research activities will be further enhanced and that China will actively participate in international science and technology organisations and large international science programmes. Toward building a collaborative relationship with Europe, by 2015, China has provided 20,000 scholarships to support Chinese students and scholars to study in European countries, and 10,000 to support EU students and scholars to study in China. Meanwhile, 2,000 Chinese students benefitted from Erasmus Mundus scholarships in 2012. In Europe, the strategic document “A Long-Term Policy for China-Europe Relations”, issued in 1995, demonstrated Europe’s intention to cooperate with China (European Commission, 1995).

As a result of cooperation support from both sides, China and Europe have kept a good track record of collaboration on research and innovation. In the first three years of the 7th Framework Programme for Research and Technological Development (FP7), Chinese researchers were the third most allocated-to recipients of funding amongst non-European researchers (European Commission, 2010). An increasing number of Chinese scholars

¹ http://eeas.europa.eu/delegations/china/press_corner/all_news/news/2015/people_en.htm

participated in the EU framework programme after the EU-China Science and Technology agreement of 1998. A number of initiatives between the EU and China also took place outside of the Framework Programmes. Consequently, the number of collaborated publications between China and the EU28 in 2014 was more than 7 times as high as that in 2000, from 2,535 to 19,241.

Apart from the general number of collaborated publications between China and the EU28, nevertheless, little is known about the collaboration patterns and strengths and weaknesses of both parties. As pointed out by Wang (2016), China's development in science is characterized by its divergent capacity across disciplines. Cooperation in the fields where China has the strongest comparative advantage is undoubtedly beneficial to its EU partners. Likewise, China is also attempting to set up collaborations in fields where Chinese researchers can learn and benefit most from its western partners.

In order to foster more profitable collaboration for both parties and to formulate options for international policy on research and innovation cooperation between the European Union and China, this paper presents an in-depth analysis of the scientific collaboration between China and the EU28, focusing on the major research priorities and benefits of these collaborations. To shed light on the initiatives of collaborations, corresponding authors are detected and classified into three categories: Chinese local, Chinese abroad, and non-Chinese.

2. DATA COLLECTION AND METHODOLOGY

The dataset was collected from Thomson Scientific's Science Citation Index Expanded (SCI-E) and Social Sciences Citation Index (SSCI). Given that the matching information of authors' names and institutes' names was provided by Thomson WoS only after 2006, we focus on the joint articles between China and the EU published between 2008 and 2014. Using our query², we obtained 51,722 papers in total, which are the joint publications between China and the EU28, in the year 2008, 2010, 2012 and 2014. These publications were downloaded and imported into SQL Server to process and analyse. In a joint publication, the corresponding author is often the one who is responsible for organizing the publication, answering queries, and communicating with co-authors. For a fund-supported publication, the corresponding author is often affiliated with the institute of the funding resource. Hence, corresponding authors are generally presumed to be the ones who took the initiatives in the project and who set up the collaborations. To this end, we in particular extracted all the information (including names, institutes and countries) for the corresponding authors. In processing the data, we first distinguish Chinese researchers from non-Chinese ones. Furthermore, by the institution names and locations, we classify Chinese researchers into two categories, namely, located in China and located in the EU. To avoid confusion, we exclude the papers published by researchers who have both Chinese and European affiliations.

The Herfindahl index (also known as Herfindahl-Hirschman Index) is applied to examine the concentration degree of collaborated fields.

² CU=China AND cu=(Austria OR Belgium OR Bulgaria OR Croatia OR Cyprus OR Czech Republic OR Denmark OR Estonia OR Finland OR France OR Germany OR Greece OR Hungary OR Ireland OR Italy OR Latvia OR Lithuania OR Luxembourg OR Malta OR Netherlands OR Poland OR Portugal OR Romania OR Slovakia OR Slovenia OR Spain OR Sweden OR England OR Scotland OR Wales OR Northern Ireland) AND PY=(2008 OR 2010 Or 2012 OR 2014).

$$H_i = \sum_{i=1}^{149} S_i^2 \quad (1)$$

Where H_i is the Herfindahl index (for a certain region or country) and S_i is the share of collaborated output in field i . A higher Herfindahl value indicates that collaborations are highly concentrated in certain fields, while a lower Herfindahl value shows that collaborations are widely distributed across different fields.

To examine the comparative advantages of China and the EU in certain academic fields, we rely on the concept of Revealed Comparative Advantage (RCA), which was proposed by Balassa (1965 and 1977). Correspondingly, the RCA's for China and EU can be calculated as follows:

$$RCA_{iCN} = \frac{PUB_{iCN} / \sum_{i=1}^{149} PUB_{iCN}}{PUB_{iW} / \sum_{i=1}^{149} PUB_{iW}} \quad i = 1 - 149 \quad (2)$$

and

$$RCA_{iEU} = \frac{PUB_{iEU} / \sum_{i=1}^{149} PUB_{iEU}}{PUB_{iW} / \sum_{i=1}^{149} PUB_{iW}} \quad i = 1 - 149 \quad (3)$$

Where RCA_{iCN} and RCA_{iEU} are the comparative advantage values for academic discipline i in China and the EU28. PUB_{iCN} and PUB_{iEU} are the publication numbers of field i in China and the EU28. PUB_{iW} is the publication number of this field in the rest of the world³. If the value of RCA (for a particular field) is higher than 1, this means that this region (either China or the EU) has a comparative advantage in terms of publication quantity in this particular academic area. Otherwise it signals a comparative disadvantage in this field.

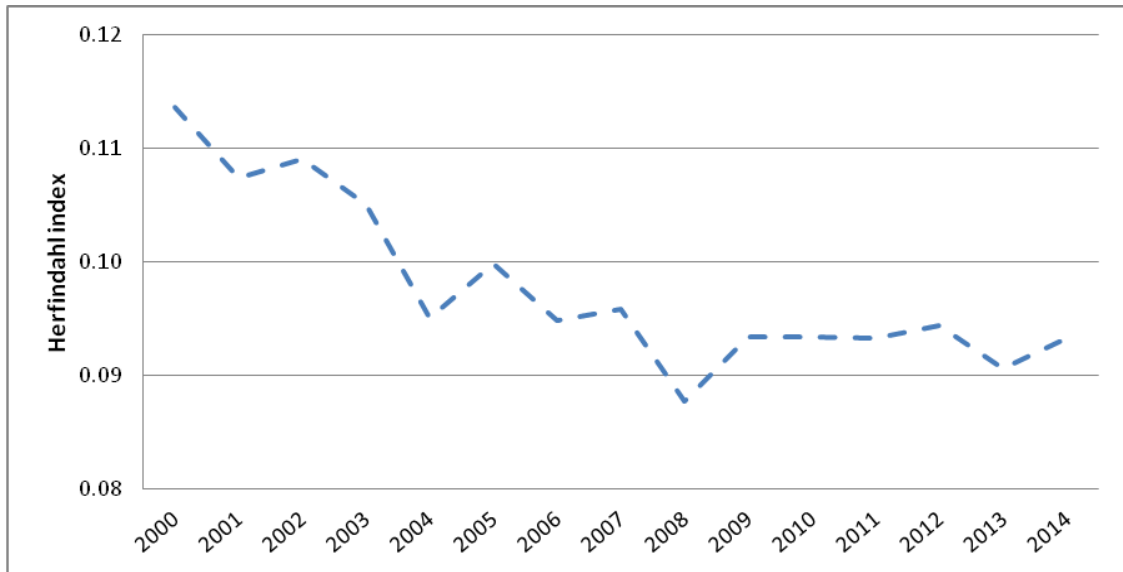
3. RESULTS

3.1 Collaboration concentration

With efforts from both sides, the jointly published academic research between China and EU28 has increased substantially, not only in the total number, but also in a wider scope. As shown in Figure 1, the Herfindal index of collaborated publications between China and the EU28 has dropped steadily over the years, which means that collaborations between China and the EU28 are becoming more and more widely distributed across different fields.

³ This is calculated by the worldwide total minus China or the EU.

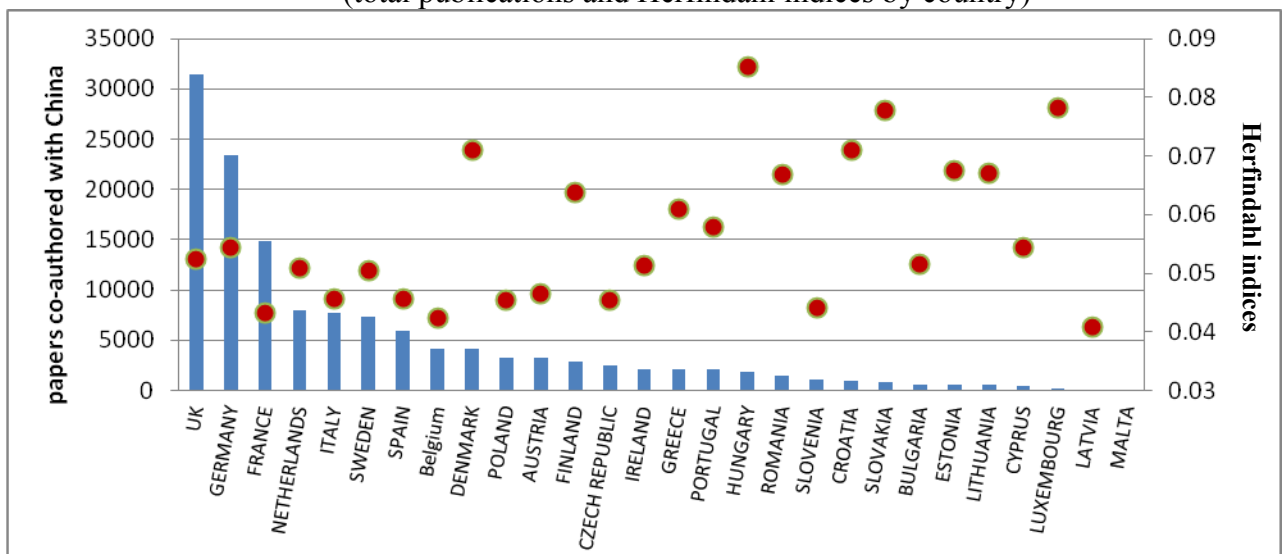
Figure 1: Herfindahl index of joint research collaborations between the EU28 and China



Source: Thomson Reuters Web of Science.

In Europe, the collaborated output came mainly from three countries (UK, Germany, and France), which together account for more than half of the total collaborated papers (with China). Six countries (UK, Germany, France, Italy, Netherlands and Sweden) accounted for 72% of the total joint output between the EU28 and China. With those countries, the collaboration has reached a wide range of research fields with a relatively low Herfindahl value (in Figure 2). However, the collaborations between China and the majority of small European countries were relatively concentrated in certain fields.

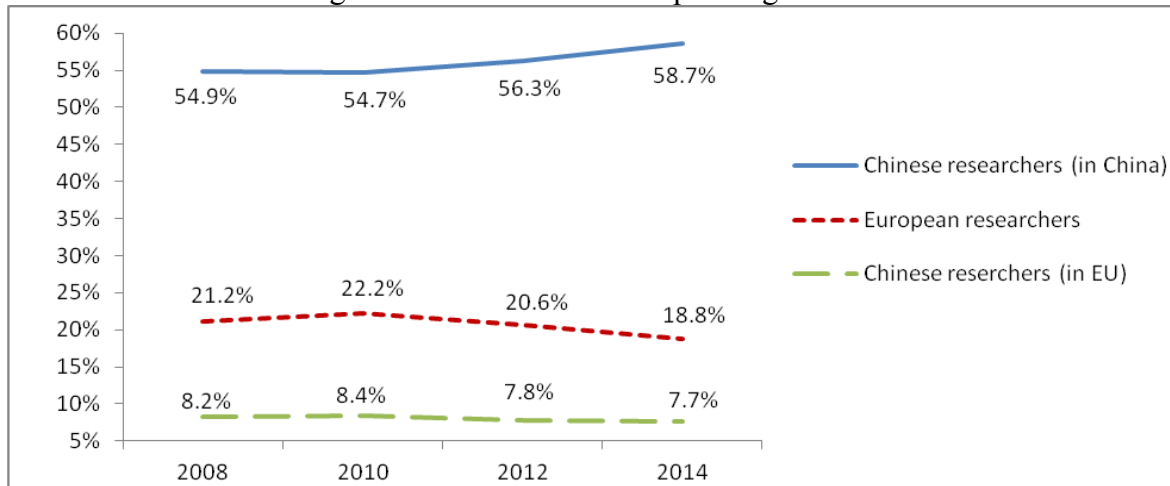
Figure 2: Research collaborations between China and the EU28 (total publications and Herfindahl indices by country)



3.2 Composition of corresponding authors

By looking at the names of authors and institutes, we distinguish three types of corresponding authors, i.e. Chinese researchers in China, Chinese researchers in the EU, and European researchers.

Figure 3: The share of corresponding authors



Chinese researchers affiliated with Chinese institutes form the major part of the corresponding authors in the jointly published academic research, accounting for more than 50 per cent (Fig.3). This proportion increased steadily after 2010 and reached 59 per cent in 2014. Non-Chinese researchers from European institutes accounted for about 20 per cent of the total corresponding authors. This share presented a decreasing trend in the later years, from 22 per cent in 2010 to 19 per cent in 2014. Chinese researchers affiliated with European institutes comprised a percentage of around 8 per cent. Compared to the percentage of non-Chinese researchers, the corresponding-author proportion of Chinese researchers in EU was considerable. The fact that a substantial portion of the corresponding authors were Chinese researchers, either working in China or in the EU, demonstrates that Chinese researchers have been playing an important role in China-EU collaboration. This finding is in line with the results of Wang, et al (2012, 2013). This implies that the academic corporation bridge between China and Europe has been mainly set up by Chinese researchers.

3.3 Fast growing fields and their RCA's

Despite the decreasing concentration in joint publication fields – which was reflected in the aforementioned Herfindahl index – some fields have appeared to be the preferential fields for EU-China joint research. Among the top 30 most collaborated fields, nine presented a growth rate of over 20% per year in terms of the EU-China collaborated publications, which are Energy fuels, Telecommunications, Science technology other topics, Instruments instrumentation, environmental sciences ecology, Oncology, Business economics, Engineering, Computer science.

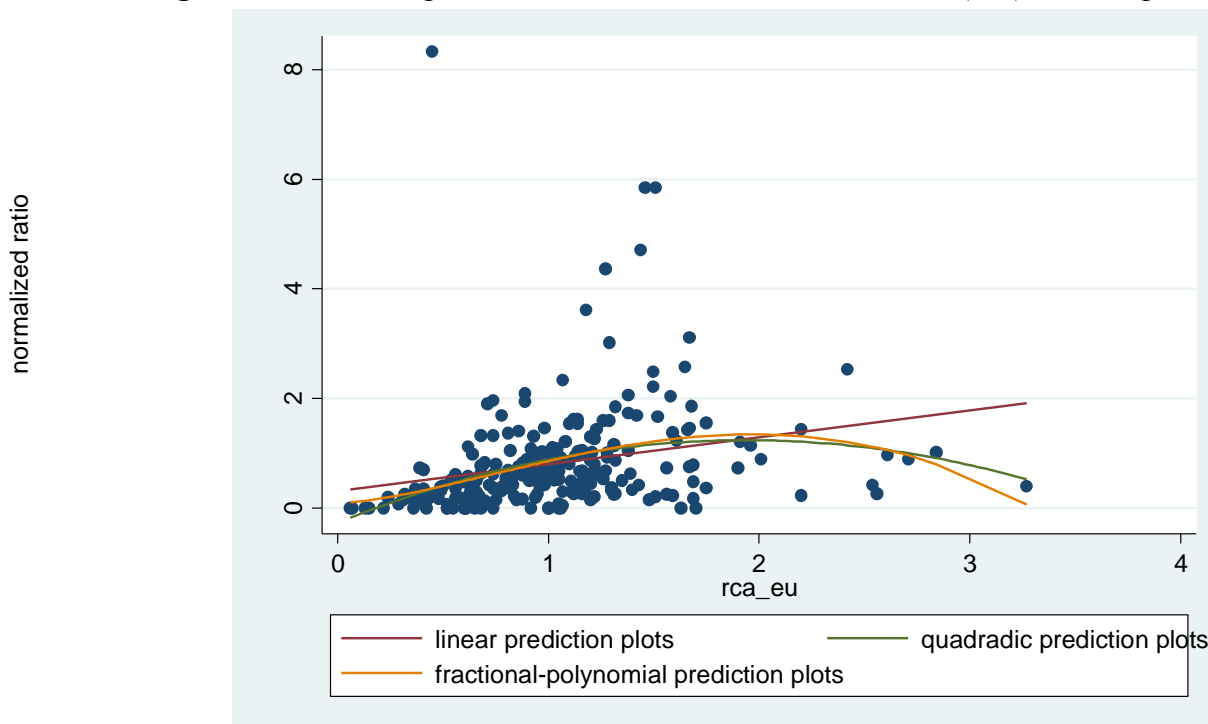
If one looks at the absolute number of joint publications between European countries and China, the top rankings no doubt go to the large countries such as the UK, Germany and France. To exclude the country-size effect, we use the following normalized ratio (NR_k^i) to calculate the relative position of joint publications in country k and field i .

$$NR_k^i = \frac{\frac{JointPub_{k,CN}^i}{JointPub_{EU,CN}^i}}{\frac{JointPub_{k,CN}}{JointPub_{EU,CN}}} \quad (4)$$

Where $JointPub_{k,CN}^i$ represents the joint publications between country k and China in field i ; $JointPub_{EU,CN}^i$ is the total joint publications between all EU countries and China in field i ; $JointPub_{k,CN}$ indicates the joint publications between country k and China in all fields; and $JointPub_{EU,CN}$ is the total joint publications between all EU countries and China in all fields. Hence, the “promising” collaborative partners, measured by the normalized ratio, are not necessarily always the most scientifically powerful countries. In the appendix, Table A1 lists the top European countries which exhibited collaboration preferences with China in the selected nine fields in the period of 2008 and 2014.

Figure 4 shows the correlation between the emerging collaborations (i.e. emerging joint publications in certain fields between European countries with China) and the comparative advantage values of the certain field in European countries (i.e. rca_eu). Most observations (242 out of 252) are located in the area where the RCA value is less than 2, where there is a positive correlation. The positive correlation indicates that more intensive collaboration is likely to happen in the field where the comparative advantage of this certain field in a particular European country is higher.

Figure 4: Relationship between normalized collaboration ratio (NR) and comparative advantage (RCA)



By analysing the specialisation of collaborations between South Africa and Germany, Schubert & Sooryamoorthy (2010) find that scientists from South Africa – as an example of a peripheral region – actively look for strong and reputable partners in the disciplines where

local specialisation is low. For the case of China, our results indicate that the joint publications with European countries were more likely from the fields where the comparative advantage (in this particular European country) is relatively high.

3.4 Dynamic RCA's

From a dynamic perspective, this section explores the changes of RCA scores in both China and European countries. Table 1 shows the changes of RCA's in both the EU and China related to the nine fastest growing collaboration fields. Except Science Technology Other Topics, all scientific fields have improved their RCA scores along with the increase of EU-China collaboration. The fact that in the fast growing collaborated fields China's RCA scores have increased relative to the EU28 indicates that China has benefited more from collaborated academic research. This links up with the fact, as shown earlier, that Chinese researchers are the dominant corresponding authors. Following the theory that corresponding authors are from the funding side and set up the collaboration bridge, it is interesting to observe that such bridges are created for joint work where China can benefit from gaining its RCA's. The European side, presumably due to not leading the joint projects, could not advance their RCA's in the fast growing collaboration fields.

Table 1: Fast growing collaborated fields between China and EU28 (ranked by growth rate)

Field	growth rate	RCA_EU				RCA_CN				EU	CN
		2008	2010	2012	2014	2008	2010	2012	2014	CHANGE (2014-2008)	CHANGE (2014-2008)
ENERGY FUELS	41.8%	0.72	0.77	0.81	0.84	1.70	1.92	1.84	1.96	0.12	0.27
TELECOMMUNICATIONS	34.3%	0.72	0.79	0.81	0.83	1.27	1.68	1.96	1.95	0.11	0.68
SCIENCE TECHNOLOGY OTHER TOPICS	29.6%	0.82	0.86	0.95	0.89	1.57	1.42	1.46	1.55	0.06	-0.02
INSTRUMENTS INSTRUMENTATION	25.1%	1.08	1.11	1.11	1.02	1.14	1.36	1.59	1.60	-0.07	0.46
ENVIRONMENTAL SCIENCES ECOLOGY	22.5%	1.10	1.12	1.15	1.16	0.95	1.09	1.06	1.04	0.06	0.08
ONCOLOGY	21.8%	1.20	1.15	1.06	0.95	0.50	0.65	0.86	1.02	-0.24	0.52
BUSINESS ECONOMICS	21.5%	1.10	1.12	1.22	1.33	0.38	0.41	0.45	0.47	0.24	0.09
ENGINEERING	20.9%	0.85	0.87	0.89	0.86	1.42	1.54	1.59	1.64	0.02	0.22
COMPUTER SCIENCE	20.9%	1.08	1.07	1.09	1.08	1.41	1.58	1.54	1.53	0.00	0.11

4. DISCUSSIONS AND CONCLUSIONS

Aiming at “revitalizing the nation through science and education strategy” (OECD 2008), China has been extending its collaborations with European countries. By extracting corresponding authors and distinguishing Chinese and non-Chinese researchers, this study provides insights into the mechanisms of joint publications between China and the EU28. Evidence shows that academic collaborations between China and the EU28 have been mainly led by Chinese researchers. In the fast-growing China-EU collaborated fields, the RCA scores in China have improved substantially. In the EU28, however, there is no such obvious improvement. This finding can be explained in two alternative ways. First, these could be the scientific fields which are strongly funded by the Chinese government, which in turn could lead to more international collaborations. Second, the collaborations with the EU28 in these fields may have helped China to strengthen these particular research fields, improving their

RCA scores. In either case, we see initial actions (or benefit) from the Chinese side, not the European side. To also benefit more from the collaboration, the EU might need to take initiatives to set up collaborations in those fields where China has higher RCA scores.

REFERENCES

Balassa, B. (1965). Trade liberalization and “revealed” comparative advantage. *The Manchester School*, Vol. 33(2), 99–123.

Balassa, B. (1977). “Revealed” comparative advantage revisited: An analysis of relative export shares of the industrial countries, 1953–1971. *The Manchester School*, Vol. 45(4), 327–344.

Bound, K., Saunders, T., Wilsdon, J. and Adams, J. (2013), *China’s absorptive state: Research, innovation and the prospects for China-UK collaboration*, Project Report. Nesta, London.

European Commission. (1995). *A long term policy for China-Europe relations*, July, COM(1995)279.

European Commission (2010), *Europe 2020 Strategy*, Brussels, 3 March 2010.

OECD. (2008). *OECD reviews of innovation policy: China*. Paris: OECD.

Schubert, T., & Sooryamoorthy, R. (2010). Can the centre-periphery model explain patterns of international scientific collaboration among threshold and industrialised countries? The case of South Africa and Germany. *Scientometrics*, 83(1), 181–203.

Suttmeier, R. P. (2014). *Trends in U . S . -China Science and Technology Cooperation : Collaborative Knowledge Production for the Twenty-First Century?* by Research Report Prepared on Behalf of the U . S . -China Economic and Security Review Commission. Retrieved

Wang, L. (2016). The structure and comparative advantages of China’s scientific research: quantitative and qualitative perspectives. *Scientometrics*, 106(1), 435–452. doi:10.1007/s11192-015-1650-2

Wang, X., Xu, S., Liu, D., & Liang, Y. (2012). The role of Chinese–American scientists in China–US scientific collaboration: A study in nanotechnology. *Scientometrics*, 91(3), 737–749.

Wang, X., Xu, S., Wang, Z., Peng, L., & Wang, C. (2013). International scientific collaboration of China: Collaborating countries, institutions and individuals. *Scientometrics*, 95(3), 885–894.



Appendix

Table A1: fast growing joint publications with China in selected fields

ENERGY FUELS				TELECOMMUNICATIONS				SCIENCE TECHNOLOGY OTHER TOPICS			
country	nr of joint publications with China	% of total EU	normalized ratio	country	nr of joint publications with China	% of total EU	normalized ratio	country	nr of joint publications with China	% of total EU	normalized ratio
SWEDEN	219	13.2%	1.62	LUXEMBOURG	8	0.5%	3.11	LATVIA	8	0.1%	1.96
DENMARK	123	7.4%	1.60	GREECE	91	6.0%	2.53	SWEDEN	636	10.4%	1.27
UK	649	39.3%	1.12	UK	711	46.7%	1.33	DENMARK	348	5.7%	1.22
FINLAND	48	2.9%	0.92	FINLAND	59	3.9%	1.23	GERMANY	1721	28.1%	1.08
FRANCE	167	10.1%	0.61	IRELAND	42	2.8%	1.16	SLOVENIA	58	0.9%	0.83
INSTRUMENTS INSTRUMENTATION				ENVIRONMENTAL SCIENCES ECOLOGY				ONCOLOGY			
country	nr of joint publications with China	% of total EU	normalized ratio	country	nr of joint publications with China	% of total EU	normalized ratio	country	nr of joint publications with China	% of total EU	normalized ratio
BULGARIA	48	4.1%	5.84	NETHERLANDS	559	14.3%	1.60	MALTA	4	0.2%	8.32
CYPRUS	34	2.9%	5.84	DENMARK	251	6.4%	1.38	LATVIA	3	0.2%	2.33
LITHUANIA	33	2.8%	4.70	IRELAND	112	2.9%	1.21	ITALY	335	17.4%	2.04
ESTONIA	34	2.9%	4.37	FINLAND	141	3.6%	1.14	SPAIN	247	12.8%	1.94
CROATIA	43	3.7%	3.61	AUSTRIA	146	3.7%	1.04	POLAND	133	6.9%	1.90
BUSINESS ECONOMICS				ENGINEERING				COMPUTER SCIENCE			
country	nr of joint publications with China	% of total EU	normalized ratio	country	nr of joint publications with China	% of total EU	normalized ratio	country	nr of joint publications with China	% of total EU	normalized ratio
UK	895	51.3%	1.46	UK	5631	46.7%	1.33	LUXEMBOURG	11	0.2%	1.44
NETHERLANDS	221	12.7%	1.42	LUXEMBOURG	21	0.2%	1.03	UK	2088	46.3%	1.32
LUXEMBOURG	3	0.2%	1.02	DENMARK	441	3.7%	0.78	GREECE	129	2.9%	1.21
BELGIUM	82	4.7%	1.01	FRANCE	1516	12.6%	0.76	FINLAND	131	2.9%	0.92
LITHUANIA	10	0.6%	0.96	SWEDEN	687	5.7%	0.69	FRANCE	646	14.3%	0.87

Note: See equation (4) for the calculation of “normalized ratio”.



A case study about the Colombian Observatory of Science and Technology: between context relevant and internationally comparable indicators

Mónica Salazar and Gustavo Crespi

Competitiveness and Innovation Division
Inter-American Development Bank

ABSTRACT

In very few countries around the world there are observatories for the design and production of science, technology and innovation (STI) indicators; mainly government organizations in charge of STI funding or national statistical agencies are the ones in charge of these activities. The purpose of this study is to assess the evolution and the results of the Colombian Observatory of Science and Technology (OCyT by its name in Spanish). The main objective is to explore if the Colombian model for the production of STI has been successful, if it is sustainable and what key institutional factors should be taken into consideration previously to replicate this experience in other contexts.

The OCyT was created because there was a clear need, mainly within Colciencias' policy makers and managers, to have regular and reliable indicators supporting the design and evaluation of STI policies. Even if Colciencias has this responsibility, it chose to create an independent institution and innovated regarding the organizational arrangement not a subsidiary agency.

As a not for profit organization and a public private partnership certainly the sustainability of the Observatory relies on the one hand, on the support provided by its associates (mainly Colciencias) and, on the other hand, on project funding. The issue at stake is if the OCyT should remain mainly as an indicator producer (its main goal and what it is paid for) or evolves into a research organization (knowledge producer), a consultant (service provider), or a combination of all.

How to maintain a balance between national needs and international standards? Indicators should help policy makers to design, monitor and evaluate policies, programs and instruments, and in that sense indicators should be customized to their necessities. In addition, indicators are also used to make international comparisons, therefore the need to adopt international standards and guidelines.

INTRODUCTION

In very few countries around the world there are observatories for the design and production of science, technology and innovation (STI) indicators; mainly government organizations in charge of STI funding or national statistical agencies are the ones in charge of these activities. The purpose of this study is to assess the evolution and the results of the Colombian Observatory of Science and Technology (OCyT by its name in Spanish). The main objective is to explore if the Colombian model for the production of STI has been successful, if it is sustainable and what key institutional factors should be taken into consideration previously to replicate this experience in other contexts.

This paper is the result of the personal experience by one of the authors of this paper who was for eight years the director of the OCyT (Salazar 2009, 2010), and recalling critically its history since its

creation and up to date. It will be complemented by a short survey to be applied to OCyT's main stakeholders.

The paper is organized as follows. The first section addresses the models that were considered when the observatory was being designed. The second section, presents a brief description of what the Colombian observatory actually does. In the third section, the results of the survey applied to the stakeholders will be presented. Finally, an analysis of the results will be made pointing to the threats, opportunities and risks that the OCyT faces.

THREE MODELS

Back in 1995, three institutional models were analyzed for the production of STI indicators (Barré, 1997), according to five functions to be performed - analysis of user needs and conception of indicators; development of methodologies and techniques; data collection and indicators production; storage and maintenance of indicators; and interpretation and use - and whether they should be concentrated in one organization or distributed in several. These models are known as: integrated functions, externalized (i.e. outsourcing) via contracting or externalized via institutionalization.

The first model proposes that the governmental science and technology (S&T) funding agency is also responsible for indicators' production, such as the National Science Foundation in the US, the Ministry of Education and Science of Portugal, Conacyt in México, or the National Agency for Research and Innovation in Uruguay. Basically in this case, demand, production, integration, use and diffusion are centralized or vertically integrated. Some of the advantages of this model are that there are good complementarities between the different tasks, accumulation of know-how, and permanent balance between supply and demand. However some of the disadvantages are lesser independence, which is a reputational asset, and coordination of various functions are not easy.

The second model proposes that the S&T agency contract the production of indicators, such as in Belgium and the Netherlands that rely on universities (e.g. NOWT). The main advantages of this model are that indicators are produced independently and that the contract regulates the needs and demands of the contractor and the supply of the producer. Some of the disadvantages are that there is a risk of non-continuity and consistency if the contract is not maintained over time.

There is a variation between the first and second model (not defined by Barré), that is when the national statistical office is responsible for STI indicator production, such as Stat Canada, INE in Spain and most European countries. Sometimes it takes up this responsibility under request from the S&T agency that contracts the production of indicators. The main advantages of this model are that the statistical office has experience on methodologies and techniques, data collection, production, storage and maintenance of indicators, and can guarantee confidentiality, usually backed by national legislation.

The third model proposes the creation of an independent organization in charge of producing the indicators, however close to the S&T public agency and funded by government, such as the *Observatoire des Sciences et des Techniques* in France, today part of the High Council for the Evaluation of Research and Higher Education (HCERES), and the Institute for Studies in Research and Education (NIFU) in Norway. In this case, conception, development of methodologies, production, and diffusion are concentrated in the new organization. This model implies that the collection of data is done by someone else, therefore a new function arises that is, the integration of information from various sources. Some of the advantages are that it can respond to various actors and stakeholders, not only the S&T agency, can guarantee continuity and quality, and can generate credibility because of technical independence.

More recently Lepori, Barré and Filliatreau (2008) have proposed three different models for the production of S&T indicators: i) the vertically integrated model, performing all five functions (same as above); ii) the data driven model, mainly related with bibliometrics and patents indicators, highly dependent on data sources; and iii) the customer-driven model. Complementing the above, Lepori, Reale, and Tijssen (2011, p 3-4) raise several issues or trends regarding the design and production of STI indicators: i) the need for performance indicators used for evaluation processes, and its political

implications; ii) the shift from the development of general purpose indicators towards customized indicators closely related to the specificities of their usage context; iii) the conceptual shift from a 'linear' process where indicators proceed from design towards (standardized) production and interpretation towards a more interactive process and iv) a broadening of data sources. Some of these issues will be raised in the last section.

THE COLOMBIAN CASE

The Colombian Observatory of Science and Technology was created in 1999, as a not-for-profit organization, promoted and led by Colciencias, the national agency that supports and funds the development of STI. Since its origin the OCyT was conceived as a public-private partnership; back then 25 organizations joined Colciencias in creating the Observatory, including universities, other public institutions, R&D centres, scientific and industry associations, among others. At present OCyT has 37 associates¹ that contribute to the financial sustainability of the organization (core funding), and act as counterbalance to any eventual Colciencias' intention to control the Observatory.

The OCyT was created because there was a clear need, mainly within Colciencias' policy makers and managers, to have regular and reliable indicators supporting the design and evaluation of STI policies (Colciencias, 1996; Ordoñez, 2002). Even if Colciencias has this responsibility, it chose to create an independent institution (the third model), and innovated regarding the organizational arrangement (i.e. a public-private partnership), not a subsidiary agency.

Two big questions can be posed when thinking about the development of the institution. How has OCyT built reliability? Why is it recognized as a producer of high quality S&T indicators? The quality of indicators produced depends highly on the quality of the data, the methodologies used, consistency, transparency and technical rigor (Salazar & Colorado, 2010). In that sense the Observatory has worked in several areas, such as: i) developing its own methodologies; ii) using and helping to improve public official data bases and administrative records; iii) supporting the elaboration of RICYT (regional) manuals²; iv) adopting and adapting OECD and RICYT guidelines and manuals; and v) creating endogenous capabilities through human capital capacity building and know-how accumulation for the production and interpretation of indicators, a key aspect in certain areas such as bibliometric indicators (see Leiden Manifesto³).

Taking into account that the Observatory generally does not collect primary data –except for S&T expenditures and for public perception of S&T- the access to third parties information is crucial, and the opportunity for data access data must be assured. The integration and treatment of the information are not a mechanical work, although usually these activities are not seen and nor duly appreciated (it is a grey labor⁴) but they are key for obtaining reliable results, requiring deep knowledge on how the data is collected and registered, definitions and attributes of the variables, and design and structure of the data bases. Because of this OCyT has become one of the more qualified users of the datasets aforementioned.

Even if the main purpose of the Observatory is the regular production of STI indicators, along with this it has built other expertise. Of particular relevance is the know-how developed in the elaboration of methodologies, not only for data collection but also for evaluation and characterization of organizations, programs and policy instruments, activities that are very context specific and customized to the necessities of the contractor.

¹ See <http://ocyt.org.co/es-es/organosyequipo>.

² Bogota Manual (2000) on innovation surveys and indicators; Lisboa Manual (2009) on ICT indicators; Antigua Manual (2015) about public perception of S&T surveys, and Buenos Aires Manual (forthcoming) a guide for indicators of academic trajectories.

³ <http://www.leidenmanifesto.org/>

⁴ See for instance: <http://blogs.lse.ac.uk/impactofsocialsciences/2014/09/01/data-carpentry-skilled-craft-data-science/>

STAKEHOLDERS' SURVEY

As a work in progress, only the questions of the survey will be presented here, the results will be available by June 2016.

1. What were the reasons for creating an independent –from Colciencias- STI observatory in Colombia?
2. Do you think that the ultimate goal or the goal pursued with the creation of the OCyT has been achieved?
3. How would you describe and qualify the model adopted by Colombia for the production of indicators?
4. How has the OCyT built confidence within its stakeholders -not only associates, but all users of the indicators- over the years?
5. From your perspective, what are the main achievements of OST in its 16 years of existence?
6. What are the risks and opportunities facing the OCyT today?
7. What are the risks and opportunities that the OCyT may face in 5 years?

The stakeholders will be grouped in several categories

- OCyT's associates.
- Users: national (policy makers, researchers, journalists, politicians, etc.) and international (OECD, RICYT, IDB, WB, Unesco, etc.).
- OCyT's scientific councilors (past and current).
- OCyT's former executive directors.

FINAL REMARKS

As a not for profit organization and a public private partnership certainly the sustainability of the Observatory relies on the one hand, on the support provided by its associates (mainly Colciencias) and, on the other hand, on project funding. The issue at stake is if the OCyT should remain mainly as an indicator producer (its main goal and what it is paid for) or evolves into a research organization (knowledge producer), a consultant (service provider), or a combination of all.

How to maintain a balance between national needs and international standards? Indicators should help policy makers to design, monitor and evaluate policies, programs and instruments, and in that sense indicators should be customized to their necessities. In addition, indicators are also used to make international comparisons, therefore the need to adopt international standards and guidelines. How to attain both needs?

If one follows the proposal of Lepori and colleagues, the challenges that the Colombian Observatory faces are major moving from the production of standardized indicators for international benchmarking to a more interactive process, customized to the user need and contextualized. For doing so, the OCyT will need to deepen its research capabilities.

REFERENCES

- Barré, R. (1997). La producción de indicadores para la política de investigación e innovación: organización y contexto institucional. In H. Jaramillo & M. Albornoz (Eds.), *El universo de la medición - La perspectiva de la ciencia y la tecnología* (pp. 5-64). Bogotá: Colciencias, RICYT, Tercer Mundo Editores.
- Colciencias (Ed.). (1996). *Hacia la construcción de un observatorio de ciencia y tecnología*. Bogotá: Colciencias.
- Lepori, B., Barré, R., & Filliatreau, G. (2008). New perspectives and challenges for the design and production of S&T indicators. *Research Evaluation*, 17(1), 33-44.
- Lepori, B., Reale, E. Tijssen, R. (2011). Designing indicators for policy decisions: challenges, tensions and good practices: introduction to a special issue. *Research Evaluation*, 20(1), 3-5.

- Ordoñez, G. (2002). La experiencia colombiana en la puesta en marcha del observatorio de ciencia y tecnología -OCyT. *CDC, 19 (51)*, 83-108.
- Salazar, M. (2009). El OCyT en el SNCTI: 10 años poniendo a prueba un modelo, presented at “Seminario Taller Tendencias en la Medición de la Investigación y la Innovación”, Bogotá, Colombia, July 16-17 2009.
- Salazar, M (2010). Challenges to the Colombian S&T indicators production system”, poster presented at 3rd European Network of Indicators Designers –ENID- Conference, Paris, 3-5 March, 2010
- Salazar, M. and Colorado, L.A. (2010). La importancia de la información en la construcción de indicadores: una verdad de Perogrullo, en M. Salazar et al, *Indicadores de Ciencia y Tecnología – Colombia 2010* (pp. 135-146). Bogotá: OCyT.



Special Globelics session proposal on: Lessons learned for priority setting and indicators relevant to the impact of research programmes in Europe and Emerging Economies. An evidence-based debate between the research and the policy-shaping community

Yannis Caloghirou* and Nicholas S. Vonortas**

*Laboratory of Industrial and Energy Economics, National, Technical University of Athens Heroon Polytechniou 9 str., 157 80, Zografou Campus, Athens.

** Center for International Science and Department of Economics and Technology Policy The George Washington University, 1957 E Street, N.W., Washington DC, 20052, USA

ABSTRACT

The purpose of this session is to present a coherent set of papers offering useful insights on research priority setting processes/activities and indicators used to measure the impact of research and technology development programmes in Europe and Emerging Economies (Brazil, Chile, Peru and Russia). In particular, the first paper focuses on the research collaborative networks funded by the European Union during the past three decades and offers a comprehensive picture of science-industry collaboration in Europe by using network indicators and providing data on the characteristics and the innovative performance of young firms participating in these networks. The second paper presents three cases of non-traditional indicators for R&D funding agencies from emerging economies and aims at contributing to the discussions on the importance of employing suitable indicators that can complement classic STI indicators. The third paper seeks to provide a critical overview of the recent exercise in the evaluation of public research institutions in Russia.

The session (180 min) aims at bringing together researchers from both developed and emerging countries as well as policy makers and will be divided into two parts . The first part will be devoted in papers' presentation and the second one in papers' discussion by invited policy experts and officials.

Proposed papers

1st paper: Network indicators for studying Research Joint Ventures

“Thirty years of European Collaboration in Research and Development: Policy-driven Research Networking and the presence of new knowledge-intensive entrepreneurial ventures”.

Yannis Caloghirou, Aimilia Protogerou and Evangelos Siokas

Laboratory of Industrial and Energy Economics, National Technical University of Athens, Greece

Since their inception in 1984, Framework Programmes (FP1-FP7) have been basic pillars of European scientific and technological development, integration and cohesion by supporting all kinds of R&D in high technology sectors and promoting cross-border, interdisciplinary networking activity (Caloghirou et al, 2004; Caloghirou et al. 2003). The emerging collaborative research networks embody the added value of bringing together different types of participating entities from different countries with complementary expertise in productive Research Joint Ventures (RJVs). In this paper we examine the structure and evolution of research collaborative networks funded by the EU Framework Programmes during the last 30 years. The analysis is based on the most recent version of the STEP to RJVs database which is presently including detailed information on all collaborative cross-national research projects funded by the European Commission in FP1 to FP7. The database has been developed and maintained by the LIEE/NTUA group. EU-funded research activity has been characterized by a considerable growth in terms of participating entities and participations across FPs resulting in substantially large networks.

EU-funded collaborative projects are establishing and expanding links between diverse organizations (firms, universities, research centers, technology users etc.) which can be equated as paths for the circulation and diffusion of knowledge and eventually the joint creation of new knowledge (Protogerou et al., 2010a; Protogerou et al., 2010b). By studying at the same time country relationships and their collaboration degree, network structure and evolution though time, and central participants' characteristics and roles in the network, our analysis provides a more detailed and in-depth picture of important aspects of the longer-lasting and more sustainable contribution of EU research public policy the so-called 'behavioural additionality', i.e. fostering collaborative learning, strengthening linkages among different types of organizations engaged in the innovative process and facilitating an extensive transmission of knowledge (Protogerou et al., 2013). Therefore, our contribution goes beyond the typical appraisals of RTD expenditures that either tend to concentrate on the

additionality of public funding in terms of input (resources added into the system) and/or output (extra private and social returns created) additionality. Furthermore, we have also collected data on the characteristics and innovative performance of young firms that participate in these networks in order to better understand their exact network role (e.g. technology specialized providers, technology users, technology developers) and how their future growth and performance may be affected by their participation.

2nd Paper: STI Indicators for Emerging Economies: Experiences from Chile, Brazil and Peru

Adriana Bin, Sergio Salles-Filho, Ana Maria Carneiro, Nicholas Vonortas, Juan Ernesto Sepulveda Alonso and Paula Felicio Drummond de Castro
University of Campinas, Department of Science and Technology Policy, Brazil

STI indicators have evolved considerably in terms of diversity and standardization during the past few decades. A consistent movement towards standardization and comparability— ending up in a series of universal indicators – has been paralleled by debate on whether these indicators can capture the diversity of socioeconomic situations around the world.

We do agree with the usefulness of such universal STI indicators in making comparisons, setting benchmarks, and defining good practices. OECD publications such as the “Frascati Manuals”, the “STI Outlook” and the more recent initiative of data convergence in the Innovation Policy Platform (a partnership of the OECD and World Bank) are good examples. Recently, the STI Outlook introduced new variables in its basket of comparable indicators. For instance, new indicators of education and entrepreneurship like “top 15-year-old performers in science” and “ease of entrepreneurship index” were added in order to capture relevant features of the innovation systems. These “new” indicators proposed and applied by OECD, although shedding light in relevant subjects not commonly used in STI analyses, do not capture important characteristics of developing countries. For instance, indicators of learning, diffusion-by-imitation, and creation of capabilities should be considered and analyzed along with classic indicators of STI.

In this manuscript we present and analyze three cases of non-traditional indicators for R&D funding agencies from emerging economies. The first one refers to the Foundation for Innovation in Agriculture of Chile (FIA), the second refers to the Brazilian Innovation Agency (FINEP), and the third one relates to the Peruvian Financing Innovation in

Agriculture (FINCAGRO). In these agencies processes of evaluation have been proposed that combine traditional-OECD STI indicators with others to measure learning, formation of capabilities, know-how diffusion and catching up processes. The FIA experience is about building a system capable to capture innovation-through-learning in agriculture. The FINEP process is a broad set of traditional and non-traditional indicators built to measure outputs and outcomes of FINEP's operations in Brazil. The FINCAGRO is similar to the FIA experience; it is directed towards financing innovation within small-holders producers in Peru.

The authoring team took part in the processes of creating and implementing indicators to evaluate those three experiences. These cases will both be analyzed against each other and compared to other experiences described in the literature. The central issue is to discuss the importance of employing “catching up” indicators along with traditional standardized indicators to create a better platform to analyze STI development in particular situations. We intend to contribute to the discussions about the importance of employing suitable indicators that can be valuable to complement classic STI indicators and to generate qualified information to help policy makers designing better strategies of STI development.

3rd Paper “Use of indicators for research and policy impact evaluation: evidence from Russia”

Konstantin Fursov and Stanislav Zaichenko

**Institute for Statistical Studies and Economics of Knowledge, National Research
University Higher School of Economics, Moscow, Russia**

Evaluation of R&D performing institutions has recently become a widely disseminated practice aimed, as declared, at improving the cost efficiency of public funding for related activities. Evidence collected from these exercises can be used for re-structuring existing research networks and for re-designing national research and technology development (RTD) programmes. The aim of this paper is that of providing a critical overview of the recent Russian exercise in evaluation of public research institutions (PRIs).

Russia is an interesting case as its national S&T system is characterized by dominance of the state in terms of funds and R&D personnel assigned to PRIs (Gokhberg and Kuznetsova, 2015). Another issue is that for the past decade a series of steps have been taken to increase performance of national R&D sector. By considering long-term strategic directives and large-scale programmes our analysis provides a more detailed picture of transformations

made towards modernization of obsolete post-Soviet institutional S&T system. It is argued that the effects of these initiatives remain unclear because of underdeveloped practice of research and policy impact evaluation. The latter as noted in (OECD, 2014) are mostly oriented towards accountability and process-control, making a minimal contribution to policy learning and strategic impact assessment.

In the paper we address the recent findings from an on-going evaluation of PRIs. A list of indicators allows considering conventional inputs (R&D expenditure and personnel) together with infrastructure capabilities like research equipment and a wider range of output categories. In particular, it includes different types of publications (journal articles, conference proceedings, books and book chapters), citations and impact-factors; results of inventive work like designs, blueprints, patents and other IPRs; and financial results such as income from technology transfer or S&T services. It is shown that while this exercise has not yet been finished and its results have not been implemented for decision-making, developed evaluation framework provides opportunities both for monitoring of R&D performance and policy impact assessment.

REFERENCES

- Caloghirou, Y., Vonortas, N.S. & Ioannides, S. (Eds.). (2004). *European Collaboration in Research and Development: Business Strategies and Public Policies*. Edward Elgar.
- Caloghirou, Y., Vonortas, N.S. & Ioannides, S. (2003). Research joint ventures: A critical survey of theoretical and empirical literature”, *Journal of Economic Surveys*, 2003, 17(4): 541-70.
- Gokhberg. L. & Kuznetsova T. (2015). Russian Federation, in UNESCO Science Report 2015, UNESCO.
- OECD (2014), *Science, Technology and Industry Outlook 2014*, OECD Publishing, Paris.
- Protogerou A., Caloghirou Y. & Siokas E. (2013). Technology fusion through EU-funded research collaborative networks. *Science and Public Policy*, 40(5), 576-590.
- Protogerou A., Caloghirou Y. & Siokas E. (2010) a. The impact of EU policy-driven research networks on the diffusion and deployment of innovation at the national level: the case of Greece. *Science and Public Policy*, 37(4), 615-647.
- Protogerou A., Caloghirou Y. & Siokas E. (2010) b. Policy-driven collaborative research networks in Europe. *Economics of Innovation and New Technology*, 19(4):349-372.
- LIEE/NTUA (2016). *STEP TO RJVs Database*. Laboratory of Industrial and Energy Economics, National Technical University of Athens.



Thirty years of European Collaboration in R&D: a Search for Key Network Indicators

Yannis Caloghirou^{*}, Aimilia Protogerou^{**} and Evangelos Siokas^{***}

*y.caloghirou@ntua.gr

Laboratory of Industrial and Energy Economics (LIEE), National Technical University of Athens (NTUA),
Heroon Polytechniou 9, Zografou, 15780 (Greece)

**protoger@chemeng.ntua.gr

Laboratory of Industrial and Energy Economics (LIEE), National Technical University of Athens (NTUA),
Heroon Polytechniou 9, Zografou, 15780 (Greece)

***vagsiok@gmail.com

Laboratory of Industrial and Energy Economics (LIEE), National Technical University of Athens (NTUA),
Heroon Polytechniou 9, Zografou, 15780 (Greece)

ABSTRACT

The purpose of this session is to present a coherent set of papers offering useful insights on research priority setting processes/activities and indicators used to measure the impact of research and technology development programmes in Europe and Emerging Economies (Brazil, Chile, Peru and Russia). In particular, the first paper focuses on the research collaborative networks funded by the European Union during the past three decades and offers a comprehensive picture of science-industry collaboration in Europe by using network indicators and providing data on the characteristics and the innovative performance of young firms participating in these networks. The second paper presents three cases of non-traditional indicators for R&D funding agencies from emerging economies and aims at contributing to the discussions on the importance of employing suitable indicators that can complement classic STI indicators. The third paper seeks to provide a critical overview of the recent exercise in the evaluation of public research institutions in Russia.

The session aims at bringing together researchers from both developed and emerging countries as well as policy makers and will be divided into two parts. The first part will be devoted in papers' presentation and the second one in papers' discussion by invited policy experts and officials.

INTRODUCTION

Since their inception in 1984, Framework Programmes (FP1-FP7) have been basic pillars of European scientific and technological development, integration and cohesion by supporting all kinds of R&D in high technology sectors and promoting cross-border, interdisciplinary networking activity (Caloghirou, Vonortas & Ioannides, 2004; Caloghirou, Vonortas & Ioannides, 2003). The emerging collaborative research networks embody the added value of bringing together different types of participating entities from different countries with complementary expertise in productive Research Joint Ventures (RJVs). In this paper we examine the structure and evolution of research collaborative networks funded by the EU FPs during the last 30 years. EU-funded research activity has been characterized by a considerable growth in terms of participating entities and participations across FPs resulting in substantially large networks. In doing so we use three different types of indicators:

- a) simple descriptive indicators that capture the critical mass/scale of research projects and the characteristics of participating organizations,
- b) basic indicators on network structure evolution and actors' centrality, and
- c) indicators capturing the propensity of young participating firms for knowledge-intensive entrepreneurship.

EU-funded collaborative projects are establishing and expanding links between diverse organizations (firms, universities, research centres, technology users etc.) which can be equated as paths for the circulation and diffusion of knowledge and eventually the joint creation of new knowledge (Protogerou, Caloghirou & Siokas, 2010a; Protogerou, Caloghirou & Siokas, 2010b). By studying at the same time network structure and evolution through time, and central participants' characteristics and roles in the network, our analysis provides a more detailed and in-depth picture of important aspects of the longer-lasting and more sustainable contribution of EU research public policy the so-called 'behavioural additionality', i.e. fostering collaborative learning, strengthening linkages among different types of organizations engaged in the innovative process and facilitating an extensive transmission of knowledge (Protogerou, Caloghirou & Siokas, 2013). Therefore, our contribution goes beyond the typical appraisals of RTD expenditures that either tend to concentrate on the additionality of public funding in terms of input (resources added into the system) and/or output (extra private and social returns created) additionality. Furthermore, we have also collected data on the characteristics and innovative performance of young firms that participate in these networks in order to better understand their exact network role (e.g. technology specialized providers, technology users, technology developers) and how their future growth and performance may be affected by their participation.

METHODS AND DATA

Data

This paper's data analysis is based on the most recent version of the STEP-to-RJVs database which has been developed and maintained by the LIEE/NTUA group. The primary information source for the database construction is CORDIS, the official information service of the European Commission. It includes detailed information on all collaborative cross-national research projects funded by EU in FP1 to FP7, i.e. information on 29,434 research projects and 69,453 different organizations with 249,300 total participations covering a 30-year period. In order to provide consistent and comparable results across FPs based on actual R&D activity in different technological areas, all the mobility, training and horizontal supportive actions were excluded.

Moreover, to identify the propensity for high-potential entrepreneurship in FPs an extra dataset was constructed using four specific criteria. First, the selected firms were young companies set up between 2002 and 2007. Second, they originate from ten European countries representing different socioeconomic models. They belong to different sectors (high-tech, low-tech and knowledge-intensive business services) and, finally, they have participated at least once in RJVs in FP6 and FP7. In total 239 young firms were identified.

Methodology

Social network analysis is employed to study our research collaborative networks to shed light into their structural characteristics, the position and role of different actors, the efficiency of operations and knowledge diffusion as well as the evolution of these aspects over time. The participants in an RJV are the structural variables or nodes that allow the analysis of the specific research activities as a system. The interaction within the network context brings to

light the dynamics of exchange among participating organizations that in the networking theory is more than the sum of the individual parts.

The networks under study can be represented simply as one-mode graphs. We assume an equal role played by all partners taking part in the same project, that is, we do not assign any particular role to organizations acting as coordinators in the R&D consortia, and we disregard the direction of ties connecting pairs of organizations.

Basic indicators on network structure

Nodes are organizations participating in EU-funded research projects and edges represent the links developed between partners in common projects. To examine network structure and its evolution through time we are focusing on indicators that provide evidence on the degree of network fragmentation and the social distance among organizations. This is because knowledge flows more easily through a highly interconnected network with short paths between individual entities.

A widely-used indicator capturing network fragmentation is the size of the giant component which provides a relative index of the degree of integration attained. Therefore, a giant component including the largest part of network's nodes indicates a highly interconnected network.

There are various indicators measuring the social distance among organizations participating in a network. The average shortest path length between any two organizations in a connected graph is the characteristic path length, while the 'longest shortest path' between any pair of nodes is the graph diameter.

Networks featuring a 'small world' property exhibit high local clustering and relatively short distances between nodes and can be characterized as relatively efficient mechanisms for knowledge creation and diffusion between nodes, i.e. two key functions of R&D collaborative networks (Cowan & Jonard, 2003). To decide whether a network has the property of being a small world, the values of two parameters - clustering coefficient C and characteristic path length L - are compared with the values of the respective parameters of a completely random network. The small world property is valid when a network is much more highly clustered than a comparable random network, but the average distance among its nodes is analogous to that of a random network (Watts, 1999).

Network centrality indicators

Nodes with high centrality are the most involved in a network and hence exhibit informational benefits which may lead to increased innovative performance. We use four different but complementary centrality indicators to assess each organization's network involvement, namely degree centrality, eigenvector centrality, closeness centrality and betweenness centrality. The first three measures assess intensity of involvement (both in terms of quantity and length of connections) while the last one assesses the type of network involvement, i.e. the ability of nodes to act as bridges between otherwise disconnected network parts (Borgatti & Everett, 2006). The centrality measures were calculated for all organizations and a synthetic index has been produced by the joint rankings of organizations in terms of these four indicators.

EMPIRICAL RESULTS AND DISCUSSION

The size of research projects

The added value of EU-funded collaborative research lies, among others, in the fact that it enables the pooling of financial and knowledge resources across national borders. This allows

research projects to achieve the appropriate size or critical mass required to accomplish scientific breakthroughs. Overall, the indicators included in Table 1 suggest that as FPs evolve through time, project scale has been increasing in terms of participants, participating countries and funding. At an R&D project level, a large consortium or a large budget would in principal be associated with improved performance because of the larger and more heterogeneous pool of resources and expertise that come together for project use (Vonortas, 2009). Nevertheless, a better understanding of what ‘critical mass’ means in the context of EU-funded research projects is required. For example, more research work is essential to identify which is the ideal number of participating entities and funding per project or which is the relationship between project scale and level of output.

Table 1. The changing characteristics of research projects across FPs

	FP1	FP2	FP3	FP4	FP5	FP6	FP7
Number of projects	828	1598	2131	4778	9232	3292	7575
Number of participants	3636	12096	15398	34620	51603	44590	21999
Average number of participants per project	4.39	7.57	7.23	7.25	5.59	13.54	11.53
Average number of different countries per project	2.85	4.33	4.03	4.29	4.80	7.41	6.44
Average EU funding* per project (million €)	-	-	1.36	1.92	1.34	3.43	3.68
Average project budget *(million €)	-	-	2.18	3.55	2.40	5.37	5.78

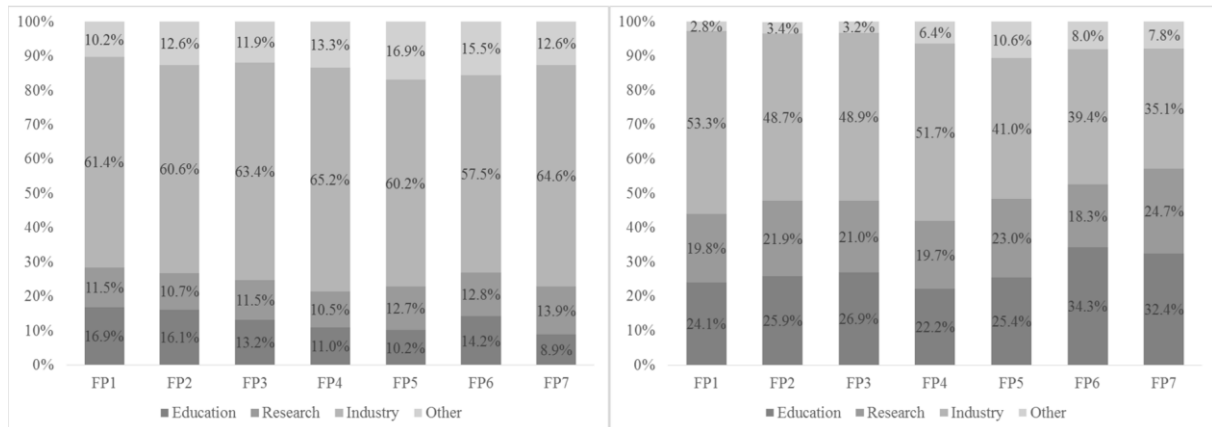
*Amounts for funding and budget are not available due to lack of the relevant data in the CORDIS database for FP1 and FP2

The characteristics of research partners

Another important category of indicators is related to the characteristics of research participants. Figure 1 illustrates both the evolution of unique participants and the evolution of their participation across FPs per organization type. The left-hand side of Figure 1 shows that the majority of participating entities in all FPs are firms.

The right part of Figure 1 points out that the participation intensity of educational and research institutions is steadily increasing (43.9% in FP1 to 57.1% in FP7), while the opposite holds for business enterprises since the relevant percentages reveal a decreasing trend across FPs (53.3% in FP1 to 35.1% in FP7). Several parameters could explain the growth of collaborative participation shares held by universities and research centres. Such a factor is the introduction of more basic-research oriented research activities (e.g. NEST in FP6) or introduction of funding instruments that are either less attractive to business firms or allow the creation of larger projects (e.g. NoEs and IPs in FP6).

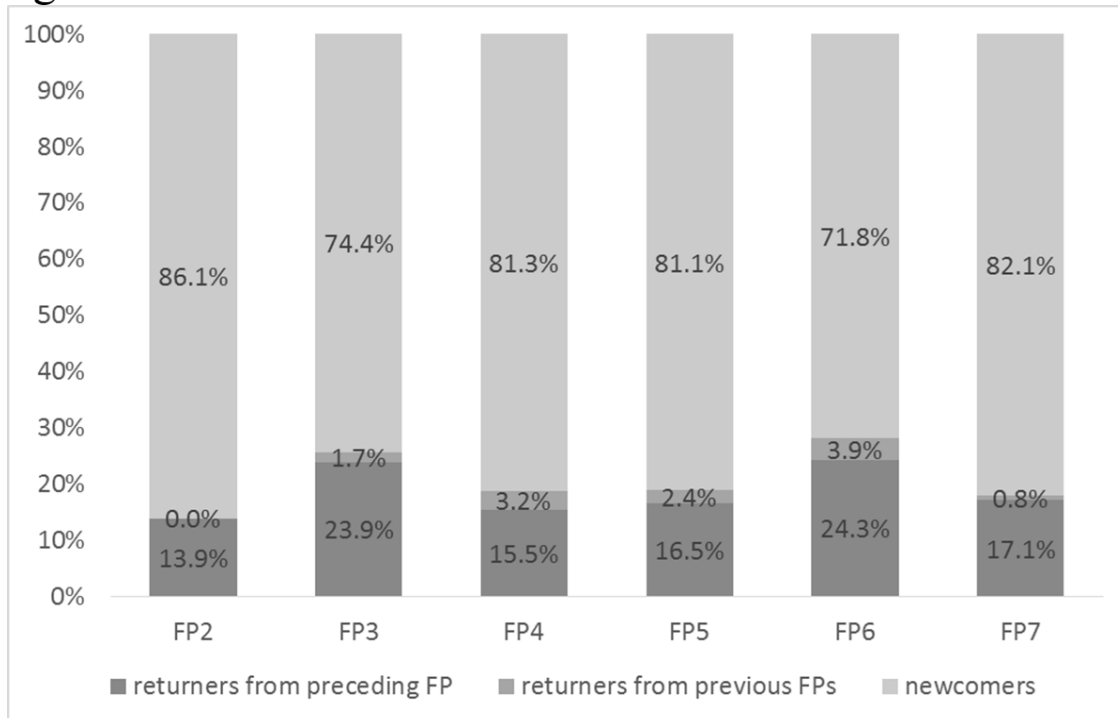
Figure 1. Unique organizations and their participation distribution among the main research actors



On the other hand the downward trend in industrial participation that has continued steadily from FP4 to FP7 may be indicating that FP activities do not come up to the overall goal of improving European competitiveness. In general, the nature of the FPs makes them more suitable research environments for academia than industry as these programmes are mainly pre-competitive. In addition, administrative complexities and bureaucracy tend to further deter industry involvement.

The organizations that overlap between successive FPs and at the same time take part in a great number of projects have the potential to create more stable relationships and thus acquire a more prominent position in the RJVs networks through time. Figure 2 indicates the returning and new actors between FP2 and FP7 highlighting that the majority of organizations participating across FPs are newcomers and mostly firms. Thus, empirical results indicate that many organizations' participation in FPs is short-lived, as 58,558 organizations (84.3%) were present in just one FP and approximately 80% of them participated in only one project as well. However, a share of participants in each FP can be characterized as "returners", i.e. organizations that have also joined either the preceding FP or earlier Frameworks. In addition, there is a relatively small number of organizations with a continuous involvement in FPs. Specifically, there is a subset of 164 actors with a stable presence across all seven FPs. Of these, the greater part is prestigious universities (53%) and research centres (27%) while firms account for 20%.

Figure 2: Distribution of returners and newcomers across FPs



An actor's involvement in successive FPs can be attributed to a number of factors. First of all it is related to the organization size and consequently their superior financial, human and organizational resources as compared to smaller-sized organizations. Furthermore, the presence of cumulative and self-reinforcing phenomena could also explain why a small number of organizations have repeat participation (Muldur, Convers, Delanghe et al., 2006). Success in obtaining funds from EU enhances the ability to raise money in the future from that (or other sources).

Networking characteristics

All networks examined are found to be tightly interconnected. The number of nodes in the largest component of the graphs representing the seven networks designate that they are highly connected and that their connectivity increases through time. Furthermore, the size of the giant component keeps increasing through time and in FP7 it covers 99.8% of all organizations present in the network. These findings highlight that the vast majority of organizations participating in EU-funded projects are, directly or indirectly, interconnected via collaboration. Therefore it can be assumed that these programmes have the potential to advance networking activity and thus foster cohesiveness and integration in the context of ERA.

Table 2. The evolution of network structural characteristics

		FP1	FP2	FP3	FP4	FP5	FP6	FP7
Nodes		1769	5144	7254	15103	19859	15953	21998
Edges		7021	50924	68018	143023	247512	399972	502185
% Edges >1		12%	15%	9%	9%	12%	11%	16%
Size of GC (%)		1535 (86.8)	4978 (96.8)	6728 (92.7)	14269 (94.5)	17730 (89.3)	15856 (99.4)	21970 (99.8)
Clustering coefficient	Actual	0.82	0.83	0.85	0.86	0.85	0.86	0.81
	Random	0.0045	0.0038	0.0026	0.0013	0.0013	0.0031	0.00208
Characteristic path length	Actual	3.84	3.15	3.29	3.28	3.08	2.71	2.73
	Random	3.61	2.86	3.03	3.27	3.08	2.47	2.62
Diameter		9	7	12	9	8	6	6

Table 2 also shows that the characteristic path length and network diameter are practically decreasing across FPs suggesting that there are better possibilities for easier and quicker information spread and knowledge diffusion through time. From the point of view of a single actor shorter distance implies easier access to the knowledge of other network actors. The RJVs networks under study exhibit “small world” characteristics, and it can be concluded that all networks examined exhibit a small world property (Table 2). Therefore, it can be assumed that the RJVs networks can be relatively efficient mechanisms for both the creation and diffusion of new technological knowledge and innovation.

Central players

Our empirical analysis so far has shown that there is a core of significant actors gaining in connectedness and significance over time by repeated participation in RJVs. These actors are usually located in strategic or central positions being those that are extensively involved in relations with other actors (Wasserman & Faust, 1994). We are next focusing on the topⁱ 1% central actors in order to shed light on the evolution of their characteristics over time.

Figure 3: Evolution of central actors' involvement in the FPs

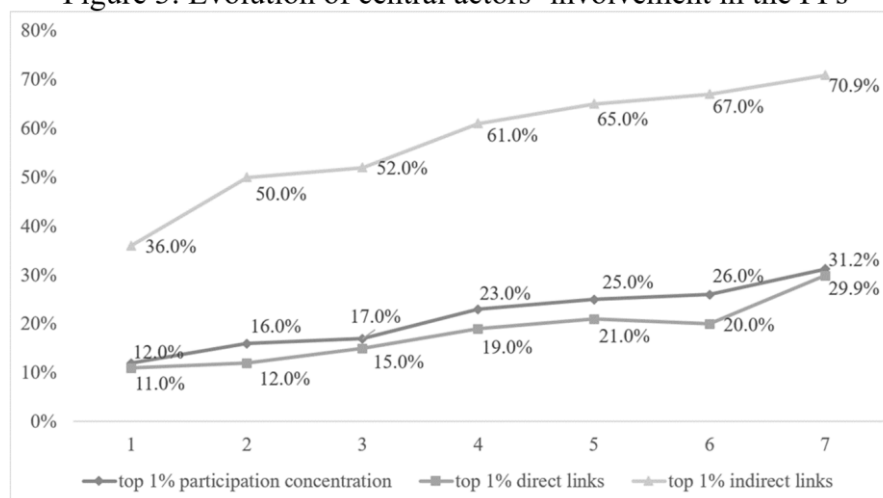


Figure 3 illustrates the evolution of participation concentration of central actors and the share of direct and indirect links they have developed within FPs in the 30 year period examined. Central actors account for a significant and increasing share of overall participations across FPs ranging from 12% in FP1 to 31% in FP7. Furthermore, the direct and indirect links that connect them to other research partners represent an increasing part of the total network connections through time. Most interestingly, their share of indirect links is quite impressive accounting for 36% of all indirect connections in FP1 and to 71% in FP7. Our findings suggest that central actors although they represent a relatively small number of nodes in each FP, are attractive to other network partners and gain in connectedness (direct and indirect) over time because they are considered as desirable partners both in terms of knowledge assets and network resources.

Table 3 illustrates the top 20 key players across the seven FPs for the time period examined (1984-2013). These organizations exhibit a stable high centrality ranking over time and they are equally represented by well-known universities and research centres whilst only one large-sized firm is included among them. In general, they join RJVs to access technological knowledge and complementary resources and skills, to promote networking and finding new partners, to share technology risk and market uncertainty or they may aim at influencing standards and technology platforms (Caloghirou, Vonortas & Ioannides, 2004).

Table 3. The top twenty central actors

Organisation Name	Type	Country	Participations	Centrality score*
Fraunhofer Gesellschaft Zur Förderung Der Angewandten Forschung EV	RES	GER	2265 (1)	5 (1)
Centre National De La Recherche Scientifique (CNRS)	RES	FRA	2064 (2)	10 (2)
Netherlands Organisation for Applied Scientific Research - TNO	RES	NET	1232 (3)	11 (3)
Consiglio Nazionale Delle Ricerche (CNR)	RES	ITA	1063 (4)	18 (5)
VTT - Technical Research Centre of Finland	RES	FIN	1021 (6)	19 (6)
Commissariat À L'Energie Atomique (CEA)	RES	FRA	1036 (5)	26 (4)
National Technical University of Athens	EDU	GRE	951 (7)	27 (14)
Consejo Superior De Investigaciones Cientificas	RES	ESP	865 (8)	35 (7)
Katholieke Universiteit Leuven	EDU	BEL	854 (9)	36 (8)
Imperial College of Science, Technology and Medicine	EDU	UK	803 (10)	42 (18)
Technical University of Denmark	EDU	DEN	660 (16)	52 (31)
Ecole Polytechnique Federale De Lausanne	EDU	SWI	671 (14)	56 (108)
Deutsches Zentrum Für Luft- und Raumfahrt EV (DLR)	RES	GER	744 (11)	61 (10)
Rheinisch-Westfälische Technische Hochschule Aachen	EDU	GER	639 (17)	62 (39)
Universität Stuttgart	EDU	GER	578 (20)	63 (46)
Lund University	EDU	SWE	570 (22)	64 (43)
Universidad Politecnica De Madrid	EDU	ESP	572 (21)	67 (43)
The Chancellor, Masters And Scholars of The University of Cambridge	EDU	UK	691 (13)	79 (24)
Kungliga Tekniska Hoegskolan	EDU	SWE	554 (23)	81 (45)
Siemens Aktiengesellschaft	IND	GER	699 (12)	81 (11)

*The centrality score is calculated for the cumulative FP1 to FP7 network

Knowledge intensive entrepreneurship in FPs

Our paper also attempts to shed some light on the potential of EU-funded research collaborative networks in fostering knowledge-intensive entrepreneurship. It does so by offering some empirical evidence on the characteristics and participation intensity of young knowledge-intensive firms in EU-funded research joint ventures, their network role, and the interaction patterns developed among them and other research actors.

Empirical results show that small and very small companies account for the largest part of young firms that participate in EU-funded research networks. In addition, young firms

established between 2002 and 2007 have a rather limited presence in EU FPs pointing out that young entrepreneurial ventures need some time to develop certain administrative and project management competences as well as the necessary research resources and technical knowledge to become attractive partners to dominant network players.

Spin-offs (45% of the young firms' sample) exhibit on average a more frequent participation in EU-funded RJVs compared to the remaining firms. Qualitative data based on information acquired from spin-offs' websites and the homepages of the research projects they are involved indicate that in general they participate in research projects that are closely related to their in-house R&D and therefore these partnerships may foster their ability of developing and launching specific commercial projects. It seems that in some cases they may enter research partnerships more easily because of their university origin or because of their founders' affiliations to certain institutions (in such cases parent universities and institutions are project coordinators).

The vast majority of organizations participating in these EU-funded projects are, directly or indirectly, interconnected via collaboration. Therefore the young firms under study are embedded in highly interconnected networks where they can have access to a large amount of technological knowledge and information held by other actors. Furthermore, newly established firms participating in EU-funded networks have the potential to develop relationships and thus exchange technological knowledge and expertise with actors exhibiting a high degree of diversity (in terms of type, sector and centrality position). Therefore, they can have access to an increased and diversified amount of resources makes EU-funded research networks suitable tools for enhancing entrepreneurial outcomes in highly competitive environments (e.g. firm performance, mergers, formation of alliances etc.). In fact, our secondary data research revealed that 38% of them hold at least one patent, they have achieved high survival rates (70% of are still in operation) while those that do not exist have in their majority been bought or merged with other firms.

Nevertheless, young firms could also be considered as attractive partners to large incumbents due to their specific technological competences and knowledge. Indeed, our secondary research indicated that in their majority these young firms are technology providers or product developers in RJVs and therefore their participation can be beneficial to the diffusion of specialized technology knowledge within these networks and in consequence contribute to the further development of EU-funded research networks.

REFERENCES

Borgatti, S.P., Everett, M.G. (2006). A graph-theoretic framework for classifying centrality measures. *Social Networks*, 28(4), 466-484.

Caloghirou, Y., Vonortas, N.S. & Ioannides, S. (Eds.). (2004). *European Collaboration in Research and Development: Business Strategies and Public Policies*. Edward Elgar.

Caloghirou, Y., Vonortas, N.S. & Ioannides, S. (2003). Research joint ventures: A critical survey of theoretical and empirical literature", *Journal of Economic Surveys*, 2003, 17(4): 541-70.

Cowan, R., Jonard, N. (2003). The dynamics of collective invention. *Journal of Economic Behavior and Organization*, 52:513-532.

Muldur U., Convers, F., Delanghe, H., Dratwa, J., Heimberger, D., Sloan, B., & Vanslebrouck, S. (2006). Building on experience: Learning lessons from past framework programmes. In *A New Deal for an Effective European Research Policy: The Design and Impacts of the 7th Framework*. The Netherlands: Springer, 93-141.

Protogerou A., Caloghirou Y. & Siokas E. (2013). Technology fusion through EU-funded research collaborative networks. *Science and Public Policy*, 40(5), 576-590.

Protogerou A., Caloghirou Y. & Siokas E. (2010) a. The impact of EU policy-driven research networks on the diffusion and deployment of innovation at the national level: the case of Greece. *Science and Public Policy*, 37(4), 615-647.

Protogerou A., Caloghirou Y. & Siokas E. (2010) b. Policy-driven collaborative research networks in Europe. *Economics of Innovation and New Technology*, 19(4):349-372.

Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

ⁱ The top 1% central actors' subgroup was chosen arbitrarily. However their removal from the RJVs networks resulted in a significant drop of the giant component initial size. We used different values than the one adopted (e.g. top 5%) to check for robustness and the main results remained unaffected.



SPECIAL SESSION PROPOSAL

PERFORMANCE INDICATORS FOR AREAS OF INNOVATION: INTERNATIONAL PERSPECTIVE

Guilherme Ary Plonski

University of São Paulo, Brazil

plonski2@usp.br

Daphne Getz

The Samuel Neaman Institute for National Policy Research, Israel

daphne@sni.technion.ac.il

Andres Barge-Gil

Complutense University of Madrid, Spain

abarge@ccee.ucm.es

Aurelia Modrego Rico

Carlos III University of Madrid, Spain

modrego@eco.uc3m.es

ABSTRACT

Introduction

The first science park (Stanford Research Park) was established in 1951. Eight years later the first business incubator (Batavia Incubator) started operations. A noticeable similitude is that neither one of those pioneering innovation niches was intentional, as they resulted from the acumen of entrepreneurial minds that perceived unconventional usages of available real estate. Science parks (*a.k.a.* research parks or technology parks, or science and technology parks) and incubators have disseminated and now operate in a large number of countries, regardless of their economic level or political ideology. Science parks and business incubators were gradually regarded as prototypical innovation habitats.

A phenomenon that has gradually surfaced since the mid 2000's is the emergence of non-traditional types of innovation niches: accelerators, catapults, innovation districts, high-tech hubs, technopoles, makerspaces, hackerspaces, co-working spaces, fab labs, tech shops, innovation labs, living labs and others. Although each of them possesses individual features, they share converging aims, which are akin to the purposes of incubators and science parks. The proliferation these models generated the need for a new and encompassing idea. The recent notion of "areas of innovation" devised by IASP, the main international trade association of science parks, headquartered in Malaga, Spain, contends for such a concept¹:

¹ <http://www.iasp.ws/the-role-of-stps-and-innovation-areas;jsessionid=46a52f94984122520b7c8a6e9b6a>. The relevance of the new concept led to a change in the name of the organization, now called "International Association of Science Parks and Areas of Innovation".



“Areas of innovation, of which science, technology and research parks (STPs) are a highly specialized type, play a key role in the economic development of their environment. Through a dynamic and innovative mix of policies, programmes, quality space and facilities and high value-added services, they:

- Stimulate and manage the flow of knowledge and technology between universities and companies;
- Facilitate the communication between companies, entrepreneurs and technicians;
- Provide environments that enhance a culture of innovation, creativity and quality;
- Focus on companies and research institutions as well as on people: the entrepreneurs and ‘knowledge workers’;
- Facilitate the creation of new businesses via incubation and spin-off mechanisms, and accelerate the growth of small and medium size companies; and
- Work in a global network that gathers many thousands of innovative companies and research institutions throughout the world, facilitating the internationalization of their resident companies.”

The global dissemination of the areas of innovation, combined with an increasing visibility of some of them generated a growing interest in assessing their performance. This interest evolved in diverse metrics, reflecting the standpoints of the different stakeholders, and the distinct frames of reference, due to the geographical, cognitive and social location of the innovation areas. Therefore, on the top of foreseen differences in measurement standards among distinct types of areas of innovation (e.g., incubators and innovation districts), dissimilar metrics are used for the very same type of areas. For example, in some countries a key measure for the performance of business incubators is the number of highly qualified jobs generated by their knowledge intensive incubatees and graduate firms, whereas in other nations the amount of capital attracted by those firms is a foremost measure.

Purpose and Intended Audience

The intended outcome of the special session is the activation of an international core group of experts and practitioners interested in performance indicators for areas of innovation. This core group could and should expand later on.

The intended audience comprises, without being limited to:

- a) Researchers and graduate students in areas such as local and regional development, entrepreneurship and innovation policy and management;
- b) Public policy makers and government officials in areas related to development planning and to science and technology, including national and regional agencies that stimulate innovation and new ventures;
- c) Managers of areas of innovation;
- d) Angel investors and professional from the venture capital community, and managers of corporate ventures;
- e) Professionals from think tanks that study and promote advanced economic development; and
- f) Organizations that develop and operate rankings.

Proposed activities

The intended form is a 90 minutes roundtable discussion, with the following preliminary activities:



1. Who is in the room? (5')
2. Perspectives from Spain, Israel and Brazil (30')
3. Glimpses from other countries present (20')
4. General discussion (25')
5. Summary of conclusions (5')
6. What's next? (5)

A concise document with the proceedings will be prepared, in order to disseminate the results of the special session and attract interested parties for further expansion of the core group.

Relevance to the Conference

The special session will benefit from areas of innovation metrics practiced in Brazil, Israel and Spain. These three countries are connected with but peripheral to the nations that are main centers of economic, scientific and technological endeavor. The organizers will also address emerging areas of innovation located in peripheral areas in each country.

Novelty

There are already substantial studies on areas of innovation indicators, mainly along two perspectives:

- ✓ Reports of actual individual practices, such as performance indicators of a specific business incubator, or KPIs established for a determined science park; and
- ✓ Broad proposals, such as national performance systems aimed at measuring the success of science parks, or to measure the economic impact of business accelerators or business incubators on a country's economy.

Nevertheless, as mentioned before, performance indicators for innovation areas is a current issue that is not yet mature and, therefore, not adequately covered by existing indicators (quantitative or qualitative).



LENGTH

90 minutes

PREFERRED NUMBER OF PARTICIPANTS

20-25

SPECIAL NEEDS

None

WARM-UP PRESENTATIONS

Daphne Getz and Eliezer Shein

Science and Technology Parks (STP) play an important role in creating a supportive Eco-System to build innovation, developing new businesses, transferring of technologies, establishing tight collaboration between academia and the industry and as a result it impacts in a positive way on the growth of knowledge and high-tech economy. It can also help in the transformation of a peripheral city into a metropolis. To successfully manage this eco-system and determine its significance, a well-defined evaluation system is needed to continually assess the performance of each STP and its influence and contribution to its harboring city.

This warm-up presentation will highlight the transformation of Be'er-Sheva - Israel's emerging high-tech hub in the Negev Region, far away from the crowded startup scene of Tel Aviv. It will propose goal-based key performance indicators of science and technology park effectiveness, benefiting from a case study of the Be'er Sheva Advanced Technologies Park (ATP).

Guilherme Ary Plonski and Désirée Moraes Zouain

In contrast with the two pioneering models in the USA, business incubators and technology parks were introduced in Brazil in the late 1980's as part of an agenda aimed at developing knowledge-based new drivers for economic and social development. Their implementation was intensely based on academic institutions, with the support of government and specific private nonprofit organizations. The nowadays more than 400 Brazilian business incubators and technology parks cooperate regionally and nationally, constituting a 'national innovative entrepreneurship movement'. A recent relevant development is the multiplication of start-ups, start-up promotion programs and accelerators. However, these new mechanisms and firms did not join the mainstream movement. In fact, a specific association of startups was created, and managed to recruit 200 members in four years, circa 5% of the estimated number of Brazilian startups. Also an association of accelerators was recently established. As a consequence, several metrics for assessing the performance have been on the run.

This warm-up presentation will highlight the transformation of Recife, in the Northeastern part of Brazil, far away from the São Paulo – Rio de Janeiro economic and technological center. It will propose key performance indicators for areas of innovation, benefitting from a case study of Porto Digital (the name means "Digital Port" in English). This endeavor was initiated in 2000 in a historic, albeit deteriorated and depressed part of the city, combining diverse types of areas of innovation in order to establish a pole of development based on world class software industry.

CHAPTER 3

Measuring Innovation



Measuring macro-level effects of the global economic recession on university-industry research cooperation¹

Joaquín M. Azagra-Caro^{*}, Robert J.W. Tijssen^{** ,***} and Alfredo Yegros-Yegros^{**}

^{*} jazagra@ingenio.upv.es

INGENIO (CSIC-UPV), Universitat Politècnica de València, Camino de Vera s/n, E-46022 Valencia, Spain

^{**} Centre for Science and Technology Studies, Leiden University, The Netherlands

^{***} DST-NRF Centre of Excellence in Scientometrics and STI Policy, Stellenbosch University, South Africa

ABSTRACT

The 2007/2008 financial crisis, and ensuing economic recession, had a direct negative effect on university-industry research cooperation in the OECD countries and other economies – it diminished the number of university-industry co-authored research publications (UICs) during the period 2008-13 by 7%. It also changed the relationship between national business expenditure on R&D and UIC output levels. Before the recession the relationship was negative, but became positive during the years 2008-2013. The few countries where business expenditure on R&D increased during recession saw UIC numbers rise. This moderating effect of the recession applies only to ‘domestic UICs’, where universities cooperated with business companies located in the same country. Micro-level research is needed to assess the contributing effects on large university-industry R&D consortia on both domestic and international collaboration patterns.

INTRODUCTION

Technological innovation and industrial R&D often relies on academic research and university-industry collaboration (Cohen et al., 2002). The financial crisis of 2008 not only ushered in an era of austerity in public finances in many advanced industrialized nations, but it also affected business sector R&D spending which fell 4.5% in 2009 when many corporate strategies shifted from long-term competitiveness to short-term survival (OECD, 2012). What was the net effect of the economic crisis and these developments, initiated both in the business sector and public sector, on the levels of university-industry R&D interaction and collaboration across the globe? There are no survey-based sources, neither at national statistical offices nor at supranational agencies such as the OECD, to systematically assess and compare trends within and across nations worldwide. To examine large-scale effects of changes in corporate R&D spending, our descriptive empirical analysis relies by necessity on a measurement approach that extracts its data from hundreds of thousands of research articles in the open scientific literature. In this paper we analyze recent trends in university-industry co-authored research publications (UICs for short), one of the main outputs of successful joint research where academics and corporate R&D staff actively collaborated to produce new knowledge. This analytical approach builds on a long tradition of UIC-based studies (Tijssen, 2012) and policy-related applications of country-level UIC data in international statistical sources such as the European Commission’s Innovation Union Scoreboard.

¹ Another version was finalist of the VII UAM-Accenture Chair Award (2015).

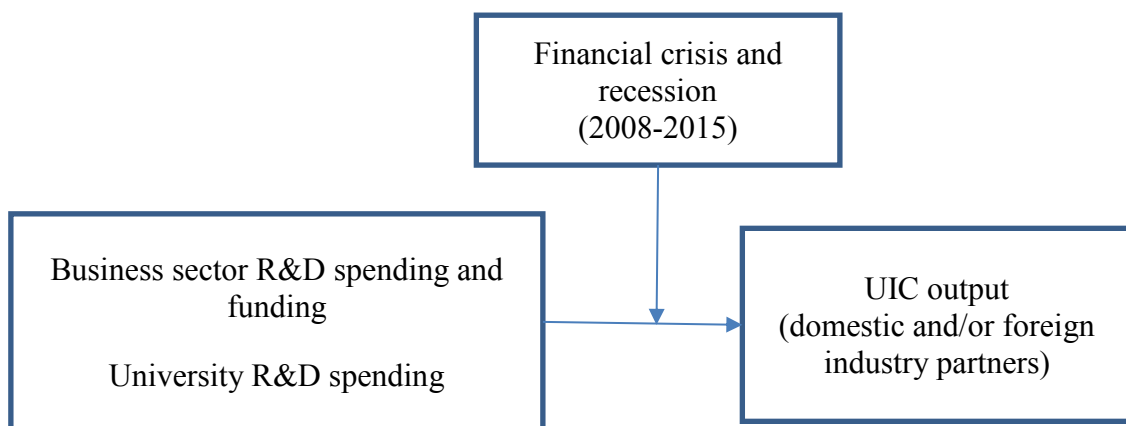
Operationalized in terms of UIC output data, two main research questions emerge:

- Did the economic crisis and recession affect UIC output levels across countries worldwide?
- Did this economic shock change the current relationship between R&D spending input and UIC output?

CONCEPTUAL FRAMEWORK

Our conceptual framework, depicted in Fig. 1, assumes a time-delayed direct relationship between R&D spending inputs and UIC output that is affected by the onset of the financial crisis and subsequent economic recession.

Fig. 1. Conceptual framework for studying the impact of the economic recession on the production of university-industry co-publications within national university systems.



THE EFFECT OF R&D ON UICS

Scientific production in the form of published articles implies underlying R&D activities at universities (Adams and Griliches, 1996; Crespi and Geuna, 2008) and business enterprises (Halperin & Chakrabarti, 1987; Chakrabarti, 1990). UICs are produced by universities as an output from successful collaboration with ‘science-based’ firms that partially outsource their scientific or technological research to universities or other specialized institutes. Typical science-based industries included pharmaceuticals and biotechnology, chemicals and food, electronics, computers and information technology. The R&D rate of return, in terms of a joint publication output such as UICs, is almost unpredictable, especially when high-risk ‘discovery’ research is involved. A wide range of models and studies, and an associated body of empirical evidence, describe the complexity and heterogeneity of linkages between business sector R&D and university research. Some studies –focusing either on transaction costs (Bruneel et al., 2010), or property rights (Rappert et al., 1999) –suggest negative relationships between those inputs and outputs. Other models suggest that knowledge absorptive capacity within companies or the open innovation paradigm (Cohen & Levinthal, 1990; Perkmann & Walsh, 2007), imply positive relationships.

HOW THE RECESSION MAY HAVE CHANGED THE RELATION BETWEEN R&D AND UICS

The economic crisis and its subsequent wave of austerity measures are likely to have had significant effects on both business sector R&D and university research. Without systematic, empirical evidence one can only speculate about the nature and extent of these impacts so far. With regards to the business sector, a downturn in available funds for in-house R&D might have sparked a heightened awareness of competitive advantages and the value of knowledge-based assets that universities may offer. This in turn could have impinged on corporate strategies and policies to either restrict the publishing research findings through UICs or otherwise (in order to protect investments and intellectual property) or, conversely, embrace resource-sharing strategies that may boost research productivity and UIC outputs. University-industry research programs can also help shape open source platforms and seeding a new R&D-based innovation ecosystems (Boudreau & Lakhani, 2009). Another advantage of such programs is that, especially in times of R&D budget constraints, companies can leverage additional funds from the public or third-sector sources. High leverage ratios can be achieved if projects are pursued in long-term consortia with established partner universities in the same country, which drastically reduces transaction costs, arising from lengthy negotiations with universities, and forces partners to relinquish IPR control rights. Following these arguments one would expect more public-private R&D collaboration but not necessarily more UICs because of the transaction costs involved in producing the publication.

As for the public sector expenditure on university research, in several countries resources have been constrained, or even cut back, because policymakers may have perceived such expenditures less likely to have immediate positive effects on job creation and economic growth in the near future (i.e. the current political cycle). Moreover, longer-term benefits of research may also accrue to other countries, thus strengthening the policy rationale for national disinvestments - either temporary or structural - in R&D subsidies for specific industrial sectors or allocation of research funds in specific fields of science. University departments effected by these austerity measures are likely seek for funding elsewhere to retain their research capacity. Those with strong pre-existing linkages to (local) R&D-intensive industries are likely to succeed in attracting such funds. The UIC production of such departments will remain at the same level. We expect to see UIC output declines at universities or countries where such favourable framework conditions are less frequent.

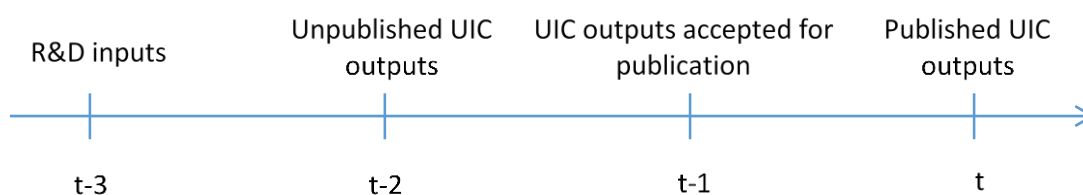
DATA AND METHODS

THE STATISTICAL MODEL

Following the knowledge production function approach (Adams and Griliches, 1996; Crespi and Geuna, 2008) we wish to fit the econometric model:

$$UIC_{it} = \beta_0 + \beta_1 R\&D_{i,t-\tau} + \epsilon_{it} \quad (1)$$

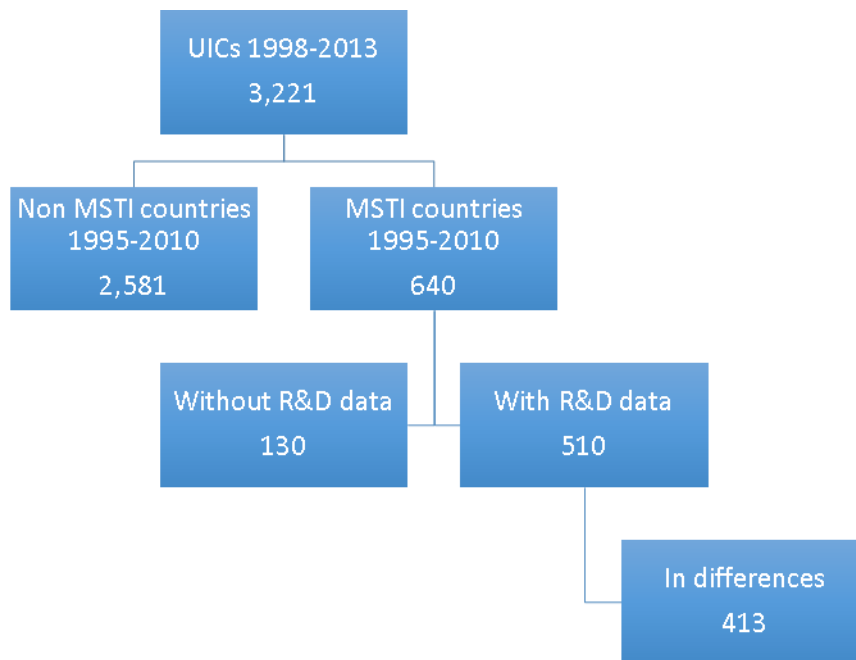
Where ‘UIC’ is the UIC frequency of country *i* in year *t*, ‘R&D’ are different indicators of R&D size, expenditure and funding of the same country τ years before *t*, and ϵ is the error term. We have set $\tau=3$ following the idea represented in Fig. 2: R&D inputs from year *t-3* generate unpublished research outputs, which are sent for peer-review in year *t-2* and will be accepted for publication in year *t-1*, and published in a WoS-indexed source year in *t*.

Fig. 2. Conceptual time lag from R&D inputs to UIC output.

VARIABLES

We measure UIC outputs through the number of university-industry co-publications. The UIC data were extracted from research publications indexed by the CWTS-licensed version of Web of Science database (WoS) which is published by Thomson Reuters. The WoS includes some 12,000 sources, i.e. peer-reviewed journals and conference proceedings, that cover the world's mainstream scientific and technical literature. Given the fact that the vast majority of industry R&D and UICs are in other fields of science, our analysis excludes all WoS-indexed research publications in the social and behavioural sciences, as well as the arts and humanities. Each UIC is assigned in full to all 'university partner countries' corresponding to the author affiliate address(es) of the university staff. Our initial dataset contains data for an incomplete panel of around 200 countries and 16 years (1998-2013). To smooth the distribution of the observations, we use the variable in logs.

We matched these data with three-year-lagged R&D statistics from OECD's Main Science and Technology Indicators (MSTI) online. This source covers 41 countries (34 OECD member states, 7 other economies), so the number of observations after matching decreases (Fig. 3 may help visualise the matching procedure). In addition, there are missing R&D data, which leaves us with a sample of 510 observations.

Fig. 3. Number of observations in the sample.

The R&D variables included are:

- A control for size: gross expenditure on R&D (GERD), at constant prices and purchasing power parities, which we take in logs.
- Ratios to measure our three predictors: business expenditure on R&D as a percentage of GDP of (BERD/GDP), higher education on R&D as a percentage of GDP (GERD/GDP) and percentage of HERD funded by industry (industrial funding/HERD).
- Plus a control for the strength of public research organisations, which in some countries is large and can act as a substitute for university R&D: government expenditure on R&D as a percentage of GDP (GOVERD/GDP).

Most variables contain non-stationary panels according to unit root tests, also when we include a trend (we conducted Im-Pesaran-Shin, Dickey-Fuller and Phillips-Perron tests). In order to avoid spurious results from regression analysis, we generate a first-difference estimator with a constant, so the model becomes:

$$\Delta UIC_{it} = \beta_1 \Delta R\&D_{i,t-3} + \delta_0 + \Delta u_{it} \quad (2)$$

Where Δ represents a year increase, the differences remove β_0 and the time constant part of the error term in equation 1, there remains an idiosyncratic error term u_{it} , and a new constant δ_0 expresses a year increase in the trend. Taking differences removes the first period from the sample and the number of observation drops from 510 to 413.

We use a dummy for identifying the recession, as in previous works (Klapper and Love, 2011; Furceri and Mourougane, 2012; Daim and Ozdemir, 2015). The dummy takes value 1 in years 2008-2013. We expand the previous model as follows:

$$\Delta UIC_{it} = \beta_1 \Delta R\&D_{i,t-3} + \gamma_1 crisis_t + \delta_0 + \Delta u_{it} \quad (3)$$

RESULTS

Table 1 presents the estimation results. Column 1 contains R&D variables only. The coefficient of the size variable, GERD, is positive and significant, which is intuitive: larger national research systems generate larger numbers of UICs. One out of our three predictors is significant, BERD intensity, with a negative sign: countries with scientifically stronger firms find universities less necessary, and fewer UICs arise.

Table 1. First-difference estimation of Δ log number of UIC outputs

	1	2	3	4	5	6
Δ Log GERD	0.51*** (0.16)	0.49*** (0.15)	0.53*** (0.14)	0.50*** (0.15)	0.50*** (0.14)	0.53*** (0.14)
Δ BERD/GDP	-0.17* (0.09)	-0.17* (0.08)	-0.22** (0.09)	-0.17* (0.09)	-0.17* (0.08)	-0.22** (0.10)
Δ HERD/GDP	-0.48 (0.50)	-0.33 (0.49)	-0.38 (0.49)	-0.35 (0.54)	-0.35 (0.49)	-0.40 (0.54)
Δ Industrial funding/HERD	0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)
Δ GOVERD/GDP	-0.42** (0.19)	-0.32 (0.21)	-0.44** (0.20)	-0.34 (0.22)	-0.30 (0.21)	-0.42* (0.21)
Recession 2008-2013		-0.06*** (0.01)	-0.07*** (0.01)	-0.06*** (0.01)	-0.06*** (0.01)	-0.07*** (0.01)
Δ BERD/GDP \times Recession 2008-2013			0.44*** (0.15)			0.43*** (0.15)
Δ HERD/GDP \times Recession 2008-2013				0.14 (0.62)		0.04 (0.59)
Δ Indl. fung./HERD \times Recn. 2008-2013					0.01 (0.01)	0.01 (0.01)
Constant	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)
Observations	413	413	413	413	413	413
Clusters	35	35	35	35	35	35
R ²	0.04	0.08	0.10	0.08	0.08	0.10
p	0.04	0.00	0.00	0.00	0.00	0.00

No multicollinearity according to VIF. Main terms centered if interacted with recession dummy. Country-clustered standard errors in parenthesis.

The coefficient of HERD intensity is not significant, suggesting that business firms lead the production of UICs, vis-à-vis universities. Industrial funding of HERD does not exert a significant influence either –a sign that numbers of UICs depend on business scientific strength rather than business attempts to orient university research.

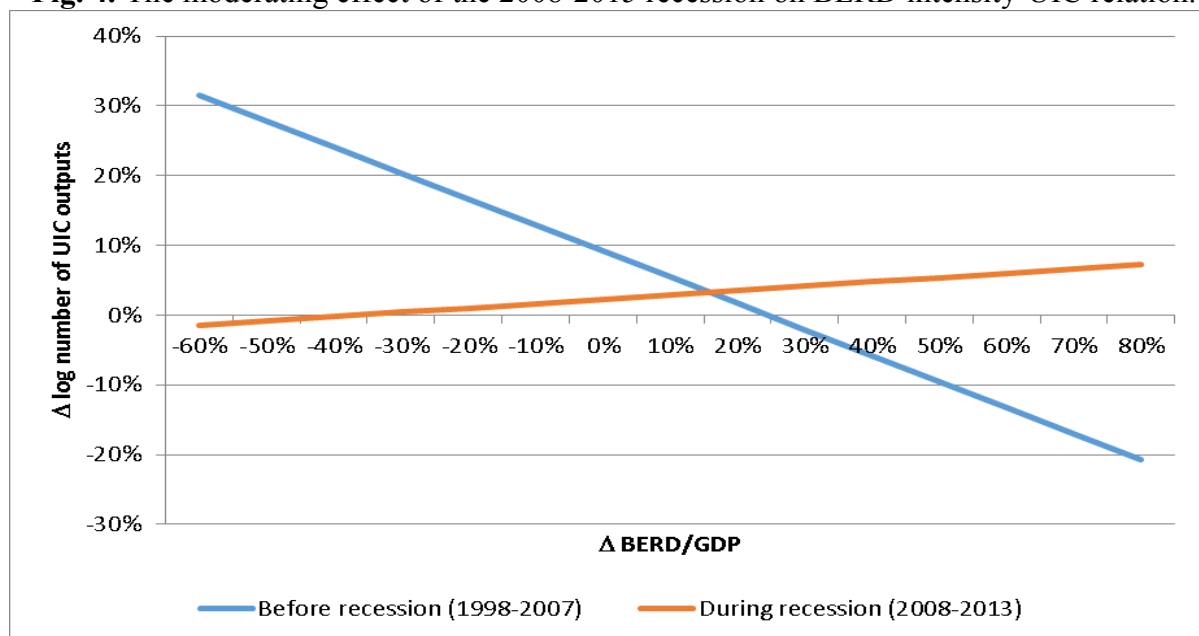
All these results are consistent across the different specifications and columns of Table 1. The impact of the control variable GOVERD intensity is somewhat erratic –always negative (indicating a substitution effect between universities and public research organisations as research partners for firms), but not always significant. Deepening into the reasons for this uneven behaviour lies beyond the scope of the paper, which focuses on university and industry, but exploratory analysis suggests that it is a case of moderated mediation with BERD intensity. In our preferred model (column 6), the coefficient of GOVERD intensity is negative, weakly significant, which we take as the most precise estimation.

In Column 2, we add the recession dummy. This is negative and significant, and so in the rest of the columns. It indicates that the direct effect of the recession on UIC outputs was an average reduction of 6-7 percent per country and year.

In Columns 3-5, we interact the recession dummy with each one of our predictors. The only significant one is the interaction term with BERD intensity (Column 3). This implies that the recession positively moderated the effect of BERD intensity on UIC outputs. The pre-recession negative impact of BERD intensity on UICs shifted to a positive impact during the recession, which suggests that UIC turned into a complementary asset to firms. In Column 6, we put all interaction terms together, and this result still holds.

Fig. 4, plotted after calculating the marginal effects of the recession at various points of BERD intensity, further illustrates the former result.

Fig. 4. The moderating effect of the 2008-2015 recession on BERD intensity-UIC relation.



The area between the two lines represents how the recession moderated the impact of BERD intensity. Decreases of BERD correspond to decreases of UIC outputs higher than the average 7%, which was the direct effect of the recession (a 0% increase in BERD intensity corresponds to the average 7%). For small increases in BERD intensity (between 0% and 40%), the effect of the recession on UICs was not significant. For large increases in BERD

intensity (from 40%), there were net gains from the recession in terms of increased UIC outputs. A closer look at the distribution of the sample shows that more than one half of the observations lie in the segment of small BERD increases, around one third in that of decreasing BERD and a bit more than ten percent in the large BERD increase segment.

The dependent variable, number of UICs, can be decomposed between domestic and foreign (68 and 32% of all UICs, respectively). We estimated the models using this breakdown. The results on the main terms are identical to the aggregate in sign and significance. The results on the interaction term change, though: they hold for domestic UICs, not for foreign ones, i.e. the moderating effect of the recession on the BERD intensity-UIC relation is only significant in the case of domestic UICs.

CONCLUSIONS

Overall, the macro-level trends with UIC production worldwide provide empirical, albeit circumstantial, evidence that the economic recession had a significant influence on research collaboration between the business sector and universities. Obviously, our country-level ‘one-size-fits-all’ explanatory model cannot grasp the real-world complex dynamics of relationships between research funding and UIC outputs. Structural determinants, supplementary to those we have now introduced and analysed, should also be investigated, notably: institutional characteristics of national science systems, the nature and scope of national research funding schemes, the effects of academic reward and incentive systems, as well as the existence of (inter)national initiatives to promote corporate R&D and university-industry cooperation.

The success of some European large public-private R&D programs during the years 2008-2014, such as the Innovative Medicines Initiative, consisting of many dedicated university-industry R&D consortia, suggests the latter strategy among pharmaceutical companies (Gunn et al., 2015). As such, the recession may have contributed to the rise of ‘open science/open innovation’ modes of cooperation (Gassman et al., 2012).

Further data gathering and detailed comparative analysis is needed to closely monitor and assess these processes and to corroborate our macro-level findings. UIC data exclude information from social sciences and humanities, whereas R&D data do not, hence refining UIC data to include social sciences and humanities would provide a more accurate match. Assessing the possible economic implications of structural changes in university-business R&D cooperation, requires micro-level case studies of the dynamic interrelationships between university research strategies, company R&D portfolios, and government R&D support initiatives.

References

- Adams, J., & Griliches, Z. (1996). Measuring science: An exploration. *Proceedings of the National Academy of Sciences*, 93(23), 12664-12670.
- Bruneel, J., D’Este, P. & Salter, A. (2010). Investigating the factors that diminish the barriers to university–industry collaboration, *Research Policy*, 39 (7), 858-868,
- Boudreau K. & Lakhani, K. (2009) How to Manage Outside Innovation, *MIT Sloan Management Review*, 50, (4), 68-76.

- Chakrabarti, A. K. (1990). Scientific output of small and medium size firms in high tech industries. *Engineering Management, IEEE Transactions on*, 37(1), 48-52.
- Cohen W & Levinthal, D. 1990. Absorptive capacity: a new perspective on learning and innovation, *Administrative Science Quarterly*, 35, 128-152
- Cohen, W., Nelson, R. & Walsh, J. (2002). Links and impacts: the influence of public research on industrial R&D. *Management Science*, 48, 1–23.
- Crespi, G. A., & Geuna, A. (2008). An empirical study of scientific production: A cross country analysis, 1981–2002. *Research Policy*, 37(4), 565-579.
- Cruz-Castro, L., & Sanz-Menéndez, L. (2015). The effects of the economic crisis on public research: Spanish budgetary policies and research organizations. *Technological Forecasting and Social Change*, forthcoming.
- Daim, T. U., & Ozdemir, D. (2015). Impact of US Economic Crises on University Research and Development Investments. *Journal of the Knowledge Economy*, 6(1), 13-27.
- Furceri, D., & Mourougane, A. (2012). The effect of financial crises on potential output: new empirical evidence from OECD countries. *Journal of Macroeconomics*, 34(3), 822-832.
- Gunn, M, Lim, M., Cross, D. & Goldman, M. (2015). Benchmarking the scientific output of the Innovative Medicines Initiative, *Nature Biotechnology*, 33, 811-812.
- Halperin, M. R., & Chakrabarti, A. K. (1987). Firm and industry characteristics influencing publications of scientists in large American companies. *R&D Management*, 17(3), 167-173.
- Klapper, L., & Love, I. (2011). The impact of the financial crisis on new firm registration. *Economics Letters*, 113(1), 1-4.
- OECD (2002). *Benchmarking Industry-Science Relationships*. Paris: Organisation for Economic Cooperation and Development.
- Perkmann, M., & Walsh, K. (2007). University–industry relationships and open innovation: Towards a research agenda. *International Journal of Management Reviews*, 9(4), 259–280.
- Rappert, B., Andrew Webster A. & Charles, D. (1999), Making sense of diversity and reluctance: academic–industrial relations and intellectual property, *Research Policy*, 28 (8), 873–890.
- Tijssen, R. (2012). Co-authored research publications and strategic analysis of public-private collaboration. *Research Evaluation*, 21, 204-215.



The impact of methodology in innovation measurement¹

Lars Wilhelmsen^{*}, Markus Bugge^{**} and Espen Solberg^{***}

^{*}lars.wilhelmsen@ssb.no

Department of industry statistics, Statistics Norway (SSB), Po.Box 8131 Dep, 0033 Oslo, Norway

^{**}markus.bugge@nifu.no

Nordic Institute for Studies in Innovation, Research and Education (NIFU), Po.Box 2815, 0680 Oslo, Norway

^{***}espen.solberg@nifu.no

Nordic Institute for Studies in Innovation, Research and Education (NIFU), Po.Box 2815, 0680 Oslo, Norway

ABSTRACT

Innovation surveys and rankings such as the Community Innovation Survey (CIS) and Innovation Union Scoreboard (IUS) have developed into influential diagnostic tools that are often used to categorize countries according to their innovation performance and to legitimise innovation policies. Although a number of ongoing processes are seeking to improve existing frameworks for measuring innovation, there are large methodological differences across countries in the way innovation is measured. This causes great uncertainty regarding a) the coherence between data from innovation surveys, b) actual innovativeness of the economy, and c) the validity of research based on innovation data. Against this background we explore empirically how different survey methods for measuring innovation affect reported innovation performance. The analysis is based on a statistical exercise comparing the results from three different methodological versions of the same survey for measuring innovation in the business enterprise sector in Norway. We find striking differences in reported innovation performance depending on how the surveys are carried out methodologically. The paper concludes that reported innovation performance is highly sensitive to and strongly conditioned by methodological context. This represents a need for increased caution and awareness around data collection and research based on innovation data, and not least in terms of aggregation of data and cross-country comparison.

BACKGROUND

Increased attention towards the importance of innovation has created a growing need for international comparisons of innovation intensity across countries. According to Smith (2005) there are three main sources that can be used to measure or proxy innovation activity. Firstly, economic indicators gathered for other purposes than innovation, but which indirectly reflect important aspects of innovation. Patent data is the most well-known example of such data, but national accounts, register data and other forms of accounting also belong to this category. Secondly, bibliometric data are often used to capture the more academic aspects of innovation activity. The third category consists of survey data, which includes both research and development (R&D) surveys and surveys dedicated to capture innovation performance.

Data on innovation performance in the private sector has been systematically collected across nations since the launch of CIS in 1992 (Gault 2013). CIS was largely based on the guidelines

¹ This work was supported by Eurostat and the Research Council of Norway

for collecting data on innovation outlined in the Oslo Manual (OECD/Eurostat 2005). As the Oslo Manual was a continuation of the Frascati Manual which dealt with the collection of R&D data, it is not surprising that the first edition of the Oslo Manual was targeting Research driven and technological innovations (Gault 2013). However, throughout the last decades the Oslo Manual has gone through three rounds of revisions with a fourth one currently in process. These revisions have in various steps expanded the notion of innovation, acknowledging to a greater extent the role of services in the economy, the importance of “soft” forms of innovation such as organizational and marketing innovation as well as a stronger focus on the characteristics of collaboration in innovation. In this sense the measurement framework is gradually improving and better reflecting current trends and actual practices in the economy and society.

However, as our study demonstrates, even small differences in survey methodology and practice may have significant effects on firms’ propensity to report innovation activity. Given the substantial differences in the way innovation surveys are implemented across countries (OECD 2013), this raises serious questions regarding the accuracy, validity and reliability of innovation data and international comparisons of innovation activity.

Innovation performance in European countries

If reported levels of innovation in CIS-data give a realistic picture of country differences in innovation activities, there must be large differences across Europe in terms of firms’ propensity to engage in innovation activity. For example, in the Community Innovation Survey for 2010-2012 (CIS 2012), the share of firms reporting any kind of innovation activity ranges from 21 per cent in Romania to 67 per cent in Germany. Within the group of Nordic countries, the share of innovative companies varies from 45 per cent in Norway to 56 per cent in Sweden. For product and process innovation the country differences are more pronounced, ranging from 31 per cent in Norway to 45 per cent in both Sweden and Finland. These country differences have also been rather consisted over time.

Such differences may of course reflect actual differences in innovativeness between countries. But they may also reflect the effects of different methods and practices in the organization and methodological accomplishment of the innovation survey. More specifically, the survey methodology may vary according to the following dimensions:

- Combined surveys vs. stand-alone innovation surveys
 - o If combined, which other survey the innovation survey is combined with; R&D-survey, ICT-survey etc.
- Whether the survey is mandatory or voluntary
 - o If mandatory, there are differences in terms of the degree of enforcement
- Response rates; these may vary from above 90 per cent to below 30 per cent
- The agency/authority who is responsible for carrying out the survey
- Types of respondents targeted / E.g. exclusion of certain sectors or small firms etc.
- Formulations and groups of questions included in the survey

Ideally, all national surveys should have a common practice regarding the dimensions above. However, in reality this is far from the case, partly due to a number of historical, practical and economic reasons.

Analytical framework and method

In order to explore whether the methodological characteristics currently used within the CIS survey framework affect reported innovation activity, a study was undertaken in conjunction with the implementation of the Norwegian CIS survey 2010 (Wilhelmsen 2012). The research questions guiding this exercise can be summarized as follows:

- *To what extent and how does the methods and practices for measuring innovation affect reported innovation performance?*
- *To what extent and how does a combined survey (i.e. covering both R&D and innovation) influence on reported innovation rates?*
- *To what extent and how does a voluntary survey influence on reported innovation rates?*

The research design chosen to investigate the two research questions was to test out three different ways to methodologically conduct the CIS survey in Norway. This implied a need to add two more samples of revised versions of the CIS questionnaire. By adding the two alternative approaches to the original survey methodology the data material constitutes three methodologies for conducting the same survey:

1. A mandatory and combined R&D and innovation survey
2. A mandatory survey only targeting innovation
3. A voluntary survey only targeting innovation

The first survey methodology was, until recently, the regular and established survey methodology in Norway consisting of a mandatory and combined survey including questions addressing both R&D and innovation. The second version was also a mandatory survey, but this survey was only addressing innovation. The third survey methodology was a voluntary survey only targeting innovation. The revised surveys were sent to two different samples, both drawn from the same population as – and not overlapping with – the sample of enterprises who received the regular combined survey. One of the extra samples receiving the revised questionnaire had mandatory reporting – with a goal of maintaining the response rate of the regular combined survey – while the remaining sample was made voluntary – and thus likely to achieve a substantially lower response rate.

Main findings from comparing the three samples

The results from the three different variations of the same survey show that the concerns raised in the research questions were valid. We find a significantly higher share of innovators in the special sample having received a survey questionnaire covering only innovations and not R&D as compared to the results from a corresponding sample from the regular, combined R&D and innovation survey. Moreover, we find that the reported innovation rates increase even further when looking at the sample where the same innovation-only survey was made voluntary. In total, the measured incidence of product and/or process innovation almost doubled going from a mandatory combined R&D and innovation survey to a voluntary innovation survey alone. The table below summarizes main results from the regular combined survey compared with the alternative mandatory innovation-only survey.

Table 1 Estimates, standard errors and t-tests; combined R&D and innovation survey vs. separate innovation-only survey

Type of innovation	N	Estimate, regular combined survey	Estimate, mandatory innovation-only survey	SE, regular combined survey	SE, mandatory innovation-only survey	Absolute t-value
<i>Product and/or process</i>	4485	868.9	1 344.3	58.3	72.6	5.11*
<i>Organisational</i>	4485	668.6	896.7	49.8	63.9	2.82*
<i>Marketing</i>	4485	707.7	878.9	54.8	62.6	2.06*

*p<.05, ∞ DF

Based on a statistical exercise comparing the results from three different ways of measuring innovation in the Norwegian business enterprise sector, the paper finds that methodological context significantly affects innovation performance. It is documented that a mandatory and combined innovation and R&D survey is likely to give the lowest innovation rates, whereas a voluntary survey only covering innovation is likely to give the highest innovation rates.

This holds true for both product and process innovators and for a combined product and/or process innovation indicator. The reported incidence of marketing and organizational innovation also increases between the regular combined survey and the mandatory extra sample, but these effects are smaller than for product and/or process innovation and only narrowly significant. We also observe a significant increase in the number of enterprises engaged in R&D activities, both internal and external. The share of innovators performing in-house R&D are only somewhat higher compared to the regular sample, but the reporting of enterprises having acquired external R&D have more than doubled.

Relationships with other explanatory variables have not been explored in depth. However, preliminary tests indicate that with the separate innovation survey, industries with a low R&D intensity observe a larger relative increase in their innovation rates compared to high-R&D industries.

Overall, our data show that there are clear and significant differences in the results depending on how the CIS survey is carried out; either separately or integrated with the business enterprise R&D survey. However, the results are not clear as to which of the resulting data sets are most accurate with respect to measuring innovation activity. Neither is it obvious that the most accurate set of results is necessarily the most useful for any particular purpose. As long as the less valid data sets are also reliable, knowledge about the different response behaviors among the enterprises may show one approach to yield advantageous properties that the alternatives may lack.

Based on the results from the different versions of the pilot survey presented above, a full-scale alternative innovation survey was carried out in Norway in 2013, covering the in-between period of 2011-13. This survey followed up the alternative 2 described above, namely a mandatory and separate survey. This exercise confirmed to a large degree the findings from the pilot exercise. As a consequence, it has been decided that the official Norwegian innovation surveys henceforth will be performed as separate mandatory surveys.

Results from the Norwegian CIS 2014 confirm again that the innovation rates from the previous pilot surveys are consistent.

Discussion of findings

By conducting three parallel survey methodologies on three comparable samples of firms within one country, we have demonstrated that differences in survey methodology may have a significant impact on the respondents' propensity to report innovation activity. Receiving a survey solely focusing on innovation as opposed to a survey first asking detailed questions on R&D performance before moving on to innovations is likely to affect the answers in many ways. Yet, little structured work has been done in terms of acknowledging sources of error occurring within the context of enterprises' response processes when it comes to business surveys (Bavdaž 2010).

Supported by a series of semi-structured cognitive interviews with enterprises carried out to better understand these results, there appears to be several interrelated factors explaining these findings. We believe that the higher innovation performance reported in the alternative surveys are at least partially caused by a larger share of respondents finding the separate innovation survey (without the R&D module) to be more applicable to how they view their own activities. As such, they may also be inclined to answer more of the remaining questions more diligently than they otherwise would have. This therefore stands in contrast to the regular mandatory and combined innovation and R&D survey. In the regular survey, which starts with a section on R&D, it is not uncommon to state early in the process of responding that: "this does not apply to me", and consequently proceed to check "no" or "not relevant" throughout the questionnaire; without giving substantial consideration to the actual questions given when the survey moved beyond R&D and on to innovation.

Another partial explanation may be that removing the R&D module from the set of questions, reduces the perceived technology and science focus of the survey, thus lowering the respondents' threshold for reporting an activity as innovative. Langhoff et. al. (2012) have hypothesized that asking for a too detailed (but non-exhaustive) breakdown of R&D activities might suppress reporting of other R&D activities that do not fit the available categories, a notion that seems to be supported by the results.

An additional possibility is that the different samples are likely to have reached different types of respondents within the targeted enterprises. We know from contact with enterprises, both for this project and in previous studies, that surveys requiring many "hard numbers" are more likely to be answered by accounting, personnel with financial oversight, or others with similar functions; many even in outsourced functions without any direct knowledge of the enterprises' activities. In this sense a separate innovation survey is more likely to be answered by someone performing a different function than the respondents in an R&D survey. Someone in a strategic management position in the enterprise would probably be more willing to generalize or to give a best guess than an accountant would be. We are also aware that some countries explicitly target the CIS to respondents in such a position (i.e. the Managing Director, Director of Operations or similar) when sending out the questionnaire.

Concluding remarks

The main conclusion we draw from this study is that methodological context matters significantly for how firms report innovation performance. While the CIS surveys in general

are comparatively well coordinated, our findings indicate that the effects of pragmatic methodological differences should not be underestimated.

Firstly, we see a need to treat the results from innovation surveys, including CIS, with caution. These results should be taken into account when comparing results from the Norwegian, or indeed any, R&D and innovation survey against results from CIS-surveys in other countries. Consequently, in order to arrange for international comparison, there is a need for a closer harmonization across countries in terms of the methods selected in comparable innovation surveys.

Secondly, the findings imply that one should treat research based on innovation data with caution. To the degree that innovation data is highly conditioned by the methodology applied in the data collection, then research using these data, in particular across countries, needs to take this into consideration.

Thirdly, the demonstrated impact of differences in survey methodology in the innovation survey signals a need to ensure greater coordination and coherence across countries in future innovation surveys. At the same time, it may indicate that similar differences may appear in related surveys such as the R&D-survey. It is generally known that there are large country differences also in the way the R&D survey is carried out. Hence, we suggest that the impact of such differences is empirically explored also for the R&D survey.

These findings also question the recommendations made by the OECD and others to combine future innovation surveys with other surveys (Arundel and Smith 2013; OECD 2010). Although combined surveys have clear advantages, such as lower costs, reduced response burden and room for linking innovation data directly with other data, we argue that such combined surveys may create unintended biases that will obfuscate the interpretation of the surveys. Furthermore, as national constraints and priorities are likely to cause countries to operate with different combinations of surveys, such issues will reduce the international comparability even further. Our conclusion is therefore that the upsides of a separate survey outweigh the downsides. We suggest that these issues are included in a broader ongoing discussion in Eurostat and OECD about future methodological best practices for innovation surveys and their measurement frameworks.

REFERENCES

- Arundel, A. and K. Smith. 2013. History of the Community Innovation Survey. In *Handbook of Innovation Indicators and Measurement*, ed. F. Gault. Cheltenham, UK / Northampton, MA, USA: Edward Elgar.
- Bavdaž, Mojca, 2010. 'The multidimensional integral business survey response model'. *Survey methodology*. Vol. 36 no. 1. Statistics Canada.
- Gault, F. 2013. The Oslo Manual. In *Handbook of Innovation Indicators and Measurement*, ed. F. Gault, 41-59. Cheltenham, UK: Edward Elgar Publishing Limited.
- Langhoff, K.; E. Dysvik; Y. Ding; B. Storbråten; and M.A. Campechano. 2012. Kvalitetssikring av data i FoU-statistikken – utfordringer og forbedringer. In *Interne dokumenter* Statistisk sentralbyrå.
- OECD. 2010. Measuring Innovation: A New Perspective. *OECD Publishing*.
- OECD. 2013. Innovation survey metadata - Wave 2006 -2008, NESTI room document. <https://www.oecd.org/sti/inno/Metada%20Innovation%20Survey.pdf>
- OECD/Eurostat. 2005. "The Oslo Manual" Manual – Guidelines for Collecting and Interpreting Innovation Data, Third Edition. Paris- France: OECD.
- Smith, K. 2005. Measuring Innovation. In *The Oxford Handbook of Innovation*, ed. J. Fagerberg; D.C. Mowery; and R.R. Nelson. Oxford: Oxford University Press.
- Wilhelmsen, L. 2012. A question of context - Assessing the impact of a separate innovation survey and of response rate on the measurement of innovation activity in Norway. Oslo: SSB.



A Critical Assessment of the Quality and Validity of Composite Indicators of Innovation¹

Dániel Vértesy*

*daniel.vertesy@jrc.ec.europa.eu

Competence Centre for Composite Indicators and Scoreboards, DDG.01,
Joint Research Centre of the European Commission, 2749 Via E. Fermi, Ispra (VA), 21027 (Italy)

ABSTRACT

While it is generally accepted that monitoring innovation system performance requires a set of indicators, there is a constant debate on whether and how composite indices can be used to summarize them. This paper enters this discussion by assessing the validity and quality of the most commonly used composite indicators of innovation.

In our framework, the validity of an index relates to the link between component indicators or aggregates and to the aspect(s) of national systems of innovation they seek to measure, while the quality of an indicator relates to its statistical properties.

To better understand validity, we discuss how the evolution of the national system of innovation concept and its use in policy has shifted demand from an advocacy to more analytical functions of composite indicators of innovation. We next examine selected composite indicators of innovation (the WIPO-INSEAD's Global Innovation Indicator, the Summary Innovation Index and Innovation Output Indicator of the European Commission and the Fraunhofer Innovation Index) in different contexts of external and internal validity and conduct global sensitivity analyses on them.

Our policy-relevant findings highlight the need for analytically stronger composites of a more limited set of indicators. We also found significant quality differences across the indices, as some included components which explain little or none of the variance in composite scores, and were more sensitive to modeling choices. The indices studied differed in how validly they represented various innovation system functions and types of innovation, and showed information relevant for a broader or a more limited set of stakeholders. We argue that further development of innovation indicators should put more emphasis on identifying tradeoffs within innovation policy, and unintended consequences of innovative activities.

INTRODUCTION

Policy makers and business strategists monitor multiple indicators to compare and benchmark the innovative performance of companies or the functioning of national innovation systems, as both the process and the outcomes of innovation activities are complex and variegated. Composite indicators have therefore been widely used to measure "innovation". Annually published country rankings for the WIPO-INSEAD Global Innovation Index or the European Commission's Summary Innovation Index keep attracting broad public attention. Over the last decade, the underlying concepts, selection of indicators, and modeling choices (weighting, aggregation methods) for these indices have been critically assessed (Grupp and Mogege, 2004; Schibany and Streicher, 2008; Grupp and Schubert, 2010; Gault, 2013, among others) but also gradually refined (Sajeva et al, 2005; Hollanders and van Cruysen, 2008; Saisana and Filippas, 2013). While composite indicators are apparently here to stay, the very

¹ This work was supported by the Innova Measure II grant, by DG-RTD of the European Commission

fact that there is an ongoing critical discussion (Barré, 2010) serves to improve their quality and ensure their validity as key stakeholders' values and interests evolve.

This paper aims to support this discussion by assessing the quality and validity of the most commonly used composite indicators of innovation. In our proposed framework, the quality of an indicator relates to its statistical properties (Saltelli, 2007; Saisana et al, 2005, 2011), while the validity of an index relates to the link between component indicators or aggregates and the conceptual framework (JRC-OECD, 2008; Saltelli et al, 2013). In the case of country-level innovation indicators, validity depends on how well an innovation index quantifies the state and evolution of what is referred to as a national system of innovation. We argue that these two elements are not absolute concepts. In its analytical part, this paper tries to answer the following two questions: (1) How valid measures are the most commonly used composite indicators of innovation? and (2) How coherent are the most commonly used composite indicators of innovation statistically?

METHODS

Our analysis rests on a qualitative and a quantitative pillar. By means of appreciative theorizing, we propose a framework of analysis for indicator validity. Rather than absolute, validity is seen as specific to a context (geographical, historical, related to a certain scientific or technological paradigm), and reflects an agreement between developers and users, which may be bound to change (Boulangier, 2007; Turnhout et al, 2007).

Using this framework, we study the link between the diffusion of the systems of innovation approach and the ever-greater interest in composite indicators of innovation. The complex, systemic interactions between the various components of inventions, research, technical change, learning and innovation (Soete et al, 2010; Freeman, 1987; Lundvall, 1992; Nelson, 1993; Edquist, 1997) calls for a set of indicators, rather than stand-alone ones to support evidence-based policies. Composite indicators summarize complex, multi-dimensional information, and make perceptible trends or phenomena not immediately detectable by single indicators (JRC-OECD, 2008). They offer an appealing tool not only to map the complexities of financing mechanisms, multi-level interactions, and the heterogeneity of innovative outcomes, but also to compromise between the partially overlapping aims and interests of different stakeholders. The advocacy function (increasing public awareness of a phenomenon) of composite indicators (Saltelli, 2007) may explain part of the success in giving the innovation systems approach prominence in the policy discourse. At the same time, limitations of both are well known: scores are driven by correlations, and cannot reveal causal links, but serve more as maps of systems or signals for system failures.

We examine selected composite indicators of innovation (the WIPO-INSEAD's Global Innovation Indicator, the Summary Innovation Index and Innovation Output Indicator of the European Commission and the Fraunhofer Innovation Index) in different contexts of external and internal validity, to make comparative qualitative assessments. These contexts (occasionally overlapping with functions of innovation systems, see Hekkert et al, 2007) include more restricted expectations of measuring the success in producing different types of innovations, to broader ones such as achieving competitiveness of selected segments of the economy, technological catch-up, or the creation of jobs and growth to speed up recovery from the financial crisis.

We next conduct global sensitivity analyses on selected composite indicators. We begin with multivariate analyses (correlation, principal component analysis, Cronbach's alpha) on the components to identify the internal structure, and the presence of a single or more latent

dimensions. We subsequently conduct a series of simulations in order to identify the robustness of country ranking and the sensitivity of scores to changing underlying assumptions, such as weighting or exclusion of indicators (Paruolo et al, 2013). A primary aim of the analyses is to assess the statistical properties of indices, and identify quality limitations – i.e., if the data structure does not correspond to a stated conceptual framework, or the presence of components that do not contribute to the variance in composite scores.

FINDINGS AND CONCLUSIONS

We find that as innovation has become prominent on today's policy agenda in advanced economies, the advocacy function of innovation indices seems to have lost some relevance in comparison with their analytical function – that is, to assess the innovation system from different aspects based on more fine-grained policy needs. Paradoxically, the most commonly used composite indicators of innovation aggregate a large set of indicators that correspond to broad definitions of innovation and framework conditions. This comes at a cost of their quality and validity, and may lead to an increased use of scoreboards and dashboards, which often overwhelms policy makers with excess information. There appears to be a void for composite indices summarizing a more targeted, more limited set of indicators, which, if carefully constructed, efficiently reduce dimensionality and complexity. The aggregation of a more limited set of indicators (i.e. sub-indices) allows for a better articulation of normative choices involved in the selection of indicators. It also offers greater transparency concerning quality and validity assessment; two interlinked tools that are necessary 'accompanying tools' in order to strengthen to support the use of composite indices.

Our analyses on the quality of the composite indicators examined revealed the presence of component indicators that explain little or none of the variance in overall composite scores. Such components are only seemingly part of the framework, misleading readers. We found differences between the various composite indices in terms of how they made explicit the presence of multiple, rather than one single latent dimension. These findings suggest that greater transparency may be necessary.

We also found that the different indicators showed a different degree of validity in the contexts examined. Many of the composites underrepresented certain types of innovation, and showed information relevant for only some of the stakeholders. This is at odds in particular with the high expectation from innovation to be a panacea. At best it nurtures false hopes; at worst policy makers may overlook unintended consequences of certain innovative activities they promote. We argue that further development of innovation indicators should put more emphasis on identifying tradeoffs within innovation policy, i.e. recognizing the down-sides of innovative outcomes (such as greater automation). From a technical point of view, more explicit use of sub-indices may serve the purpose of supporting the analysis capacity, even if those meet the needs of fewer stakeholders.

REFERENCES

- Barré, R. 2010. “Towards socially robust S&T indicators: indicators as debatable devices, enabling collective learning”. *Research Evaluation*, 19(3), pp.227-231.
- Boulanger, P-M. 2007. “Political uses of social indicators: overview and application to sustainable development indicators”. *Int. J. Sustainable Development*, Vol. 10, p. 14-32.
- Edquist, C. 1997. *Systems of innovation: technologies, institutions and organizations*. London; Washington, D.C.: Pinter.
- Freeman, Christopher. 1987. *Technology, policy, and economic performance: lessons from Japan*. London; New York: Pinter Publishers
- Gault, F. 2013. “Innovation indicators and measurement: an overview”, in: Fred Gault, (ed.), *Handbook of Innovation Indicators and Measurement*, Edward Elgar, Cheltenham, UK and Northampton, MA, USA, pp. 3-40
- Godin, Benoit. 2004, “The obsession for competitiveness and its impact on statistics: the construction of high-technology indicators” *Research Policy* 33, Issue 8, October 2004, Pages 1217–1229
- Grupp, H. and Mogege, M. E. 2004. “Indicators for national science and technology policy: how robust are composite indicators?” *Research Policy*, 33, Issue 9, November 2004, Pages 1373–1384
- Grupp H. and T. Schubert. 2010002E “Review and new evidence on composite innovation indicators for evaluating national performance” *Research Policy* 39 (2010) 67–78
- Hekkert, M. P., Suurs, R. A., Negro, S. O., Kuhlmann, S., & Smits, R. E. 2007. Functions of innovation systems: A new approach for analysing technological change. *Technological forecasting and social change*, 74(4), 413-432.
- Hollanders, H., and van Cruysen, A. 2008. “Rethinking the European innovation scoreboard: A new methodology for 2008-2010.” *Inno-Metrics Publication*, European Commission.
- JRC-OECD. 2008. *Handbook on constructing composite indicators: methodology and user guide*. Paris: OECD.
- Lundvall, B-A. 1992. *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning*. London: Pinter.
- Paruolo, P., Saisana M. and Saltelli A. 2013. “Ratings and rankings: voodoo or science?” *Journal of the Royal Statistical Society: Series A*. 176(3), 609–634
- Saisana M., D’Hombres B., Saltelli A. 2011. “Rickety Numbers: Volatility of university rankings and policy implications.” *Research Policy* 40, 165–177

Saisana M., Saltelli A., Tarantola S. 2005. “Uncertainty and sensitivity analysis techniques as tools for the analysis and validation of composite indicators”, *Journal of the Royal Statistical Society A* 168(2), 307-323.

Sajeva, M., Gatelli, D., Tarantola, S. and Hollanders, H. 2006, “Methodology Report on the European Innovation Scoreboard 2005, European Commission, EUR 21883 EN

Saltelli, A. 2007. “Composite indicators between analysis and advocacy”. *Social Indicators Research*, 81(1), 65-77.

Saltelli, A., Guimarães Pereira, Â., Van der Sluijs, J. P., & Funtowicz, S. 2013. “What do I make of your latinorum? Sensitivity auditing of mathematical modelling”. *International Journal of Foresight and Innovation Policy*, 9(2-3-4), 213-234.

Schibany, A. and Streicher, G., 2008. The European innovation scoreboard: Drowning by numbers?. *Science and Public Policy*, 35(10), pp.717-732.

Soete, L., Verspagen, B., ter Weel, B., 2010. Chapter 27 - Systems of Innovation, in: Hall, B.H., Rosenberg, N. (Eds.), *Handbook of the Economics of Innovation, Vol.2, Handbook of the Economics of Innovation*. North-Holland, pp. 1159–1180.

Turnhout, E., Hisschemöller, M., & Eijsackers, H. 2007. “Ecological indicators: between the two fires of science and policy”. *Ecological indicators*, 7(2007), 215-228.



INNOVATION STRATEGIES IN LATIN AMERICAN FIRMS¹

Fernando Vargas *

*vargas@merit.unu.edu

UNU-MERIT, Boschstraat 24, Maastricht, 6221 AX (The Netherlands)

ABSTRACT

During the last fifteen years, several Latin American countries have launched new policies to spur the structural transformation of their economies. In parallel, the availability of R&D and innovation statistics is greater than ever before. However, most of the new policies have been backed by the use of simple aggregated R&D and innovation indicators. This work will use a unique dataset of IS from several LA economies to produce a detailed analysis of innovation strategies of LA firms. These results will facilitate a more comprehensive comparison of innovation performance between LA countries and a richer benchmark analysis with economies from the EU. Furthermore, it will allow to measure the explanatory power of industry and country conditions in the heterogeneity of innovation strategies and to test the pertinence of the OECD sector-based technological classifications in the LA context.

During the last fifteen years, several Latin American countries have launched new policies to spur the structural transformation of their economies. In parallel, the availability of R&D and innovation statistics is greater than ever before. However, most of the new policies have been backed by the use of simple aggregated R&D and innovation indicators. This work will use a unique dataset of IS from several LA economies to produce a detailed analysis of innovation strategies of LA firms. These results will facilitate a more comprehensive comparison of innovation performance between LA countries and a richer benchmark analysis with economies from the EU. Furthermore, it will allow to measure the explanatory power of industry and country conditions in the heterogeneity of innovation strategies and to test the pertinence of the OECD sector-based technological classifications in the LA context.

¹ The author is thankful for the support of the Competitiveness and Innovation Division of the Inter-American Development Bank, which has provided that microdata needed for the realization of this project.

INTRODUCTION

Taking advantage of the called ‘commodities super-cycle’ period, some governments in the Latin American (LA) region decided to allocate fiscal resources towards strengthening or setting new initiatives in science, technology, and innovation (STI). As an example, since the mid-2000’s, countries like Chile, Colombia, Peru, and Paraguay, set new national funds aimed to promote the transition from natural resources, towards knowledge-driven economies (Crespi & Dutrénit, 2014). Technical inputs that backed the creation of these funds came mainly from comparative analysis with wealthier industrialized economies. The diagnosis was straightforward: there is a significant gap in innovation investments, especially in R&D expenditures. Therefore, most of these programs and policies are aimed to close these gaps.

The comparison of indicators was possible because of the spread of measurement activities of innovation and R&D expenditures. After pioneering experiences implementing innovation surveys (IS) by Argentina, Chile, Colombia, Mexico, Uruguay² and Venezuela, each of them with at least one survey by the end of the 90s; in the 2000s, Brazil, Costa Rica, Peru, and more recently Ecuador, El Salvador, Dominican Republic, Panama and Paraguay, developed their own attempts to measure innovation and R&D activities in the private sector. These surveys were implemented in parallel to a debate about how to measure innovation in the LA context. Some agreements in this regard, mainly addressing innovation in the manufacturing sector, were summarized in the Bogota Manual (Jaramillo, Lugones, & Salazar, 2001). Currently, the Oslo Manual provide guidelines for measuring innovation in the context of so-called “developing countries” (OECD/Eurostat, 2005). However, some discrepancies about how innovation should be measured in LA remains, which has translated into comparability issues between innovation statistics of LA countries and also with indicators based on the Community Innovation Survey (CIS) data. As a consequence, only a set of basic indicators has been used for STI benchmark analysis. Perhaps due to the absence of a regional agency empowered to coordinate innovation measurement initiatives³, the available information has not been fully exploited. Analysing patterns and distributions of firms’ innovative behaviour, going beyond the distinction between innovators and non-innovators, could increase our understanding of innovation in LA, feeding STI policy design with better inputs.

This research provides a first detailed quantitative analysis of how LA firms are innovating. Making use of a unique dataset of IS from eleven LA countries, we will empirically detect the combination of activities that describes how LA firms innovate. In addition to enriching the current discussion regarding innovation policies in LA, adding the ‘how firms innovate’ component to the ‘how much is invested in innovation’ argument, this research will also shed light on the role of industry and country conditions in the prevalence of innovation strategies. This work provides a richer description of innovation procedures followed by LA companies, permitting the comparison with similar analysis conducted using CIS data (Frenz & Lambert, 2009; Huang, Arundel, & Hollanders, 2010; Leiponen & Drejer, 2007; Srholec & Verspagen, 2012).

² The first attempt of measuring innovation in Uruguay dates back to 1988.

³ Although the Network for Science and Technology Indicators –Ibero-American and Inter-American– (RICYT) plays a crucial role in the coordination and comparison of Science and Technology indicators, the challenge of translating this convening power to the agencies in charge of innovation statistics, remains.

BACKGROUND

The differences in aggregated R&D investments between LA countries and high-income industrialized economies are well documented (Crespi, Navarro, & Zuñiga, 2010). The gap remains even after controlling by economic structure (Maloney & Rodríguez-Clare, 2007), which suggests that a more detailed analysis is needed for understanding the differences on private R&D investments. First, in any economy, R&D is not the most common way to innovate. Indeed, Gault (2010) summarized some of the stylized facts, emphasizing that most companies rely on non-R&D activities for innovation and that a significant share of the R&D expenditures is made by very few companies. Understanding if the LA problem is due to a low number of firms performing R&D, or low intensity of the R&D investments, is a critical input for public policies design.

On top of that, a comprehensive set of STI policy-mix should include support for other ways in which firms innovate. In this regard, the diagnosis for the LA region, besides the gap in total innovation investments, is less conclusive. The composition of the innovation efforts depends, to a large extent, on the resources and the technological level of the firms (Wernerfelt, 1984). Therefore, a closer look at the information at this level of analysis is critical for understanding the dynamics of innovation. One of the most visited approaches is the distinction between firms that rely on the internal development of technologies (MAKE) and those that acquire external technologies (BUY). Among others, Cassiman and Veugelers (2006), and Piga and Vivarelli (2004) found different effects and potential complementarities between the two mentioned strategies. Even more interestingly, Hou and Mohnen (2013) found that, while in low-income countries BUY strategy leads firms to productivity increases, in middle-income countries, which is the case of most LA countries, is the combination of MAKE and BUY what leads to higher economic performance. The relevance of the national context in innovation decisions at the firm level was also highlighted by Srholec (2011), although remarking that the (estimated) effect is limited.

The importance of firms' capabilities on determining the innovation strategies is also highlighted in Bayona, García-Marco, and Huerta (2001), and Veugelers and Cassiman (2005). The authors focused on 'close' and 'cooperative' innovation strategies founding that only firms that reach critical internal capacities are able to benefit from external collaborations. Other binary approaches to innovation strategies have found strong complementarities between two theoretically defined modes of innovation: codified scientific and technical knowledge (Science, Technology and Innovation (STI)), and learning-by-doing and interactions (Doing, Using and Interacting (DUI)) (Jensen, Johnson, Lorenz, & Lundvall, 2007), which highlights the potential problems of focusing policies on the STI-mode only.

Perhaps the more detailed analysis of innovation strategies has been carried out using CIS data for several countries. Huang et al. (2010) defined four types of innovation strategies, with which they were able to produce a detailed description of the aggregated innovative capacities of fifteen European countries. Their analysis remarks the heterogeneity of the innovative firms that do not invest in R&D, raising the need for better indicators for describing how firms innovate. Srholec and Verspagen (2012), using hierarchical factor analysis and clustering techniques, detected five innovation strategies prevalent in thirteen European economies. The variance of the 'ingredients' (Research, User, External, and Production) that in different combinations compose these strategies is only slightly affected by the economic sector and the country where the firm belongs. In this case, firm's particular resources and

capabilities were more relevant for defining the innovative strategy that framework conditions. The heterogeneity of innovation strategies, even within the same industry, is also highlighted by Leiponen and Drejer (2007) while studying CIS data from Finland and Denmark. Frenz and Lambert (2009) also find common patterns of modes of innovation even when comparing a larger set of countries. However, the slight relevance of sector and country features may be driven by the relatively low degree of heterogeneity⁴ among the countries analyzed by the mentioned studies, in comparison to LA economies.

This short review of the empirical evidence on innovation strategies emphasizes the need for a better description of how firms innovate to understand country-level innovation performance. This need is even more clear for LA countries, where STI policies goals are mostly defined by total innovation investment, not considering how firms innovate. Making use of a unique dataset of IS data from eleven LA countries, this research will provide a first detailed analysis of the innovation strategies in LA, which will allow a richer comparative analysis within the region and with similar exercises in the EU.

METHODOLOGY

Data

This research makes use of a merged dataset of seven LA IS. Most of the IS in LA share a common structure that contains modules about firms' general characteristics, innovation activities and expenditures, innovation output, human resources, access to finance, impacts of innovations, protection of innovations, cooperation for innovation, sources of information, and main obstacles for innovation. However, some differences between the questionnaires raise challenges for comparing the data extracted. Issues that goes from differences on how questions are phrased to the scope of questions referring to the same topic reduces the possibilities of conducting a straight comparison. Given the objective of the paper at hand, we have chosen⁵ to work with the IS waves and modules that allow for a better comparison in the areas that describe how firms innovate.⁶ Although in some countries IS covers several sectors, the majority of them are still concentrated in manufacturing. Therefore, for the sake of comparability, we restricted the sample to work with only manufacturing firms. The countries and its respective IS wave, included in this research are: Argentina (2001), Chile (2009), Dominican Republic (2010), El Salvador (2013), Panama (2008), Peru (2012), and Uruguay (2006). Table 1 describes the variables used in the analysis. Because the aim of this research is to describe the behaviour of the innovative firms, we restrict the sample to firms that invested in innovation in the period covered by the respective survey. Observations with a missing value in any of the variables of interest were also dropped from the data. The final sample size is 3,008 firms (970, Argentina; 438, Chile; 67, Dominican Republic; 281, El Salvador; 139, Panama; 753, Peru; and 360, Uruguay).

⁴ Although Frenz and Lambert (2009) included Brazil together with other eight high-income OECD countries in their analysis.

⁵ We have access to more IS waves from Argentina, Chile and Uruguay. For the other four countries we are using the currently available IS. IS from Colombia (2005) and Costa Rica (2008) were not included because of comparability issues in key variables.

⁶⁶ The detailed procedures for harmonizing these datasets will be available in a forthcoming publication (Crespi & Vargas, 2016).

Table 1: Variables included in analysis

Variables	Description
Innovation activities	
Intramural R&D	(0/1) firm has performed intramural R&D for introducing innovations
Extramural R&D	(0/1) firm has contracted external R&D for introducing innovations
Machinery, hardware and software	(0/1) firm has invested in machinery, hardware or software for introducing innovations
Acquisition of external knowledge	(0/1) firm has invested in acquisition of licenses, patents, know-how for introducing innovations
Training	(0/1) firm has engaged in training for its employees for introducing innovations
Others	(0/1) firm has performed or invested in other activities (consultancies, engineering, design, market research) for introducing innovations
Sources of information for innovation	
Internal sources	(0/1) firm used internal sources of information for their innovation activities
Market sources	(0/1) firm used clients, suppliers, competitors or consultants as sources of information for their innovation activities
Science sources	(0/1) firm used universities, research centres or government agencies as sources of information for their innovation activities
Other sources	(0/1) firm used other sources (such as conferences, internet, publications) as sources of information for their innovation activities
Innovation output	
Product innovation	(0/1) firm has introduced product innovation
Process innovation	(0/1) firm has introduced process innovation
Marketing innovation	(0/1) firm has introduced marketing innovation
Organizational innovation	(0/1) firm has introduced organizational innovation
Collaboration in innovation activities	
Collaboration breadth	(0 to 1) Index measuring collaboration with different types of organizations for innovation activities

Methodological approach

The main objective of this paper is to detect the innovation strategies performed by LA firms. The first stage of the extraction is done by performing a principal-component factor analysis (PCF) over variables that better describes firms' decisions regarding the innovation process. That is, the type of innovation activities performed, the sources of information used for these activities, collaboration patterns for innovation and the type of innovation outcome (product, process, organizational or marketing). As Table 1 shows, most of the variables used in this analysis are binary but one that is categorical⁷, therefore PCF is done by analysing the polychoric correlation matrix⁸ among the studied variables. Only factors with eigenvalues greater than 1 (Kaiser criterion) are selected. Factors are rotated by a direct oblique transformation (oblimin) that allows for correlation among factors. These extracted factors represent the main practices that firms follow for innovating.

⁷ Although the index takes values from 0 to 1, is calculated as the sum of different type of collaboration partners over the total of possible type of partners (7), therefore taken only 8 possible values.

⁸ Weighted by the expansion factor when available. The analysis was also performed with the unweighted sample and results did not change substantially.

After the PCF, a cluster k-means procedure is performed for identifying groups of firms that follow a similar combination of innovation practices, representing the innovation strategies. Finally, four cluster solution is imposed on the cluster k-means process to get a composition of firms that facilitate interpretation and comparison with similar studies in other regions.

RESULTS

Table 2 presents the results of the PCF. Four factors have been detected which represents the main practices of innovation in the sample of LA firms. Clear differences arise from the composition of each factor. We have labeled it “Product-development”, “Searching”, “Adopting knowledge”, and “Process modernizing”.

Product-development: This factor is mainly defined by the relevance of R&D activities, both intra and extramural, the acquisition of disembodied technologies and a focus on product innovation. The fact that both intra and extramural R&D are common in this factor, together with the acquisition of disembodied knowledge is in line with the complementarities between the MAKE and BUY approach found in the literature. Besides the focus on product innovation, the importance of the other innovation activities and marketing innovation suggests that the placing into the market of the newly products developed is a critical complement of the innovation approach. The relevance of R&D in this factor may make it similar to the “research-oriented” innovation ingredients found in Leiponen & Drejer (2007) and Srholec & Verspagen (2012), or the New-to-market innovating factor, from Frenz & Lambert (2009). However, it differs strongly in regards the lack of relevance of the scientific sources of information and the high importance of the acquisition of external knowledge. Both may be a consequence of the lack of complementary knowledge in the local scientific system for the R&D performers.

Searching: This factor encompasses the practice of relying on information from different external sources, more intensively market and “other” sources, but also scientific sources, for nurturing the innovation process. This externally-oriented practice does not involve investments in the acquisition of external technologies nor working in partnerships for innovation, which can be expected in firms that lack critical capacities to absorb technologies or learn from their innovation partners. Furthermore, no any focus on innovation outputs is observed but slightly towards product innovation. Altogether, this factor is not similar to any of those found in the previous analysis in European countries.

Table 2: Results of Principal-Component Factor analysis

Variables	Factor 1: Product development	Factor 2: Searching	Factor 3: Adopting knowledge	Factor 4: Process modernizing
Intramural R&D	0.87	0.09	0.07	-0.05
Extramural R&D	0.53	0.13	0.43	-0.21
Machinery, hardware and software	-0.01	0.03	-0.03	0.91
Acquisition of external knowledge	0.46	-0.09	0.42	-0.05
Training	0.17	-0.49	0.57	-0.16
Others	0.47	-0.31	0.36	-0.06
Internal sources	-0.10	0.14	0.65	0.14
Market sources	0.00	0.85	0.18	0.03
Science sources	-0.12	0.51	0.60	0.24
Other sources	0.14	0.92	-0.04	-0.12
Cooperation breadth	0.28	0.18	0.42	0.28
Product innovation	0.89	0.04	-0.20	0.16
Process innovation	0.12	-0.20	0.18	0.67
Marketing innovation	0.47	-0.27	-0.05	0.45
Organizational innovation	0.13	-0.52	0.29	0.10
Proportion of variance explained by each factor	0.23	0.18	0.17	0.13

Adopting knowledge: This practice is mostly defined by the critical role of the internal sources of information together with the exploitation of external knowledge. Is interpreted as firms looking at the knowledge market to acquire what they need to implement their internally sourced innovation ideas. This factor also shows a high factor score of the collaboration breadth. The external focus of these activities seems to affect the organizational structure of the firm, observed through the coefficient of organizational innovation which together with the relative relevance of process innovation, we interpret as an efficiency-oriented practice. This factor is similar to the one named “external” in Srholec & Verspagen (2012), but in the case of LA firms, the exploitation of external knowledge is based on disembodied technologies and “soft” practices such as training.

Process modernizing: This practice is in line with what is one of the most common narratives about innovation in LA, which is the investment in machinery as a source of innovation. Indeed, this factor is mainly characterized by the acquisition of embodied technology and a strong focus on process innovation. This practice is also found in Frenz & Lambert (2009). Therefore we label this factor with the same name.

To what extent the prevalence of this innovation practices depends on the sector or country conditions? Following Bell & Pavitt (1992), Castellacci (2008), or Pavitt (1984) and the relevance of sectoral and country conditions for technological regimes, we should expect a high level of relevance of both in the composition of innovation practices. On the other hand, the above-reviewed evidence from CIS data shows the intrinsic firm-dependent definition of approaches to innovation. We test which of these views holds while explaining the variance of each of the factors detected in the previous stages through an ANOVA-type III model.

Table 3: Percentage of the variance of each factor explained by sector, country and its interaction (ANOVA, Type III)

Factor	Country	Sector	Sector Country	Firm
Product development	1.5%	2.6%	3.8%	92.1%
Searching	35.9%	0.5%	2.0%	61.5%
Adopting knowledge	0.6%	0.9%	3.6%	94.9%
Process modernizing	1.7%	1.2%	3.5%	93.6%

The analysis presented in Table 3 shows that most of the variance is explained at the firm-level, which is in line with what has been found in Europe (Frenz & Lambert, 2009; Leiponen & Drejer, 2007; Srholec & Verspagen, 2012). However, it is worth noticing that variance of the less knowledge-based approach to innovation is remarkably explained by the country-level.

Table 4 shows the results of the cluster analysis, presenting the mean of the factor scores of each detected innovation practice. It can be clearly seen that there is a group of firms that perform high in the “Product development,” “Adopting knowledge”, and “Process modernizing” factors, which are those following a “high-profile” innovation strategy. The second group of firms is mostly performing a “process modernizing” approach with a slight relevance of “product development”, labeled here as a “Production oriented” strategy. In a third level, the group of firms that mainly relies on a combination of external sources of technology with a focus on efficiency gains, named here as “Adopters”. Finally, the group of firms mostly defined by the importance of the “Searching” factor, show limited investments in externally sourced technologies and some importance of product development, which is labeled as “Imitators”.

Table 4 Cluster Analysis: Innovation strategies

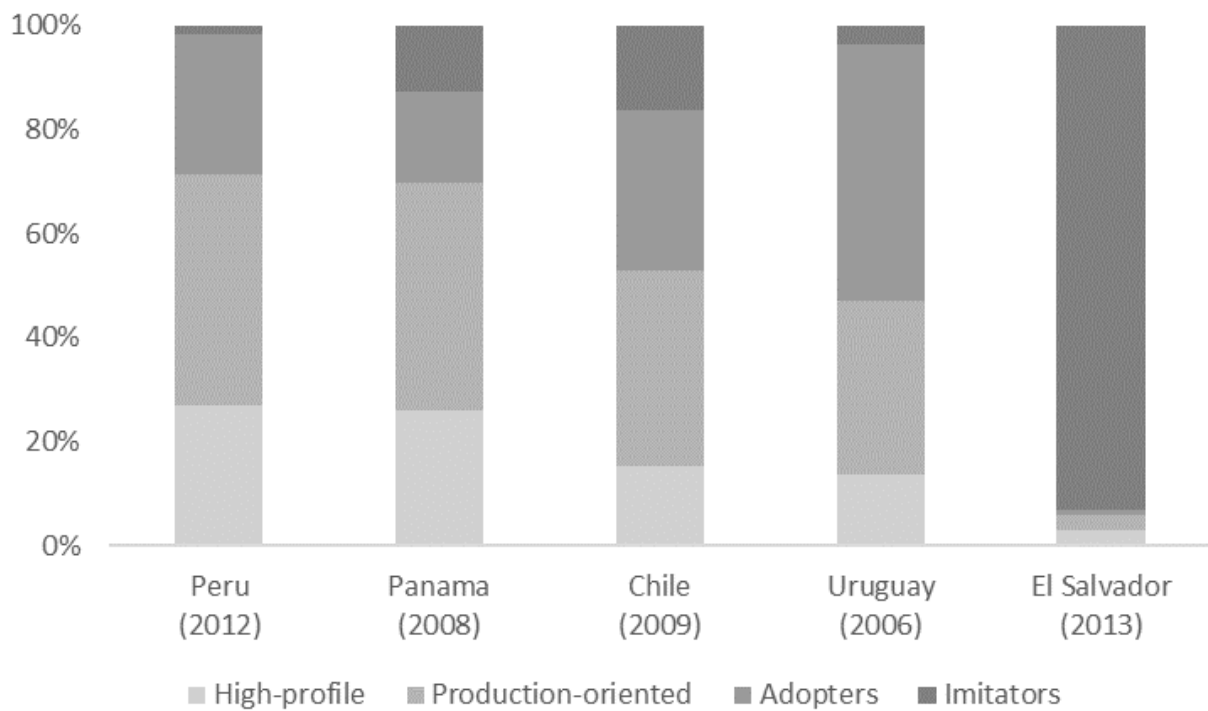
Clusters: Strategies	Factors			
	Product development	Searching	Adopting knowledge	Process modernizing
High-profile	0.90	-0.31	0.65	0.93
Production oriented	0.52	-0.10	0.04	1.01
Adopters	0.05	-0.17	0.49	0.61
Imitators	0.44	0.53	0.25	0.52

Is there any dominant innovation strategy in the sample of countries analysed? An exploratory approach is presented in Figure 1, where the share of each innovation strategy is presented for Chile, El Salvador, Panama, Peru, and Uruguay.⁹ From here it can be seen that Chile, Panama, Peru and Uruguay have a non-trivial share of firms innovating through “High-profile,” “Production-oriented”, and “Adopters’ strategies. Furthermore, the “Imitation” strategy is less common in the countries above and, on the contrary, is strikingly predominant in El Salvador.

⁹ These are the countries in the sample that provide appropriate expansion factors (inverse of the sampling fraction).

Certainly, the differences between El Salvador and the rest of the countries is correlated with the relevance of country conditions for explaining the “Searching” approach to innovation. On the other hand, the size of the firm may be a good instrument for estimating specific firm capabilities that can explain the prevalence of innovation strategies. Indeed, most of the empirical evidence of innovation in LA put the size of the company as one of the main determinants of innovation decisions but not necessarily the intensity of those investments (Crespi, Tacsir, & Vargas, 2016; Crespi & Zuniga, 2012). Figure 2 shows the prevalence of innovation strategies according to the size of the firm, defined by the number of employees. It can be seen that a larger share of large innovative firms engages in “High-profile” innovation strategies, in comparison with medium and small firms. Although, and somehow unexpectedly, this same segment of firms has the higher relative share of “Imitators”.

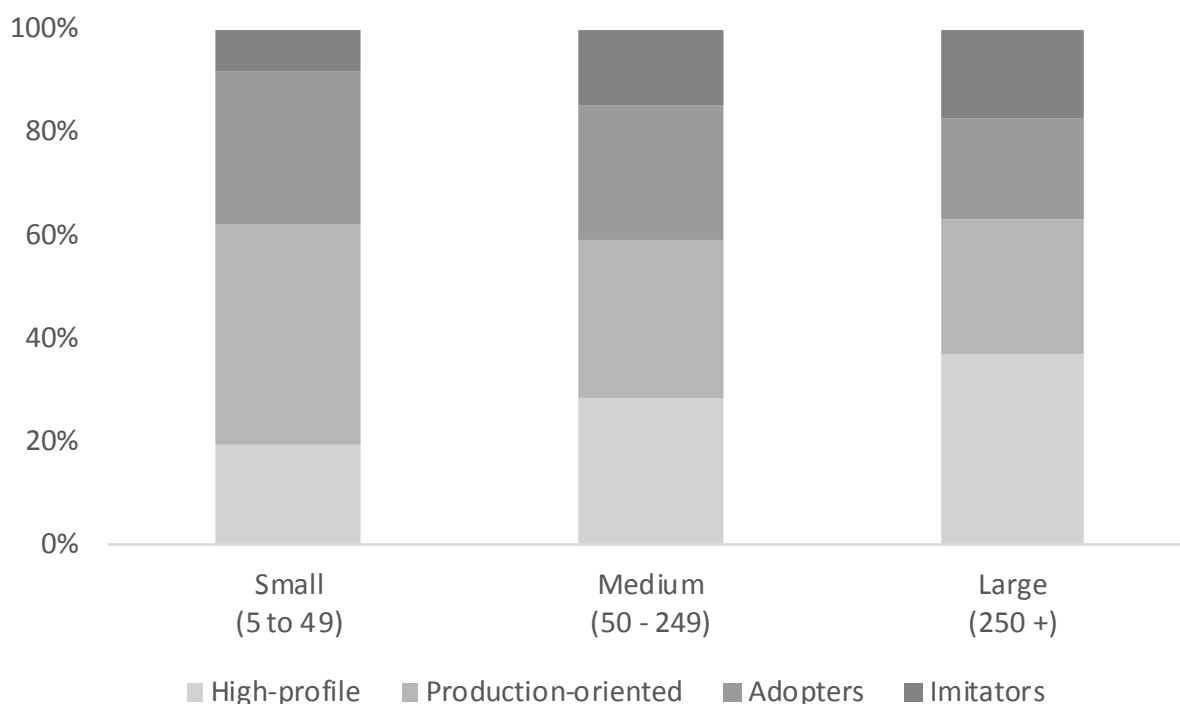
Figure 1: Distribution of innovation strategies in selected LA countries



CONCLUDING REMARKS

In the current scenario of renewed interest in industrial and innovation policies in LA countries, to increase our understanding about how firms innovate in this context is crucial for effective policy design. The paper at hand makes use of a unique dataset of IS from seven LA economies to produce a detailed analysis of most common innovation strategies in LA. These results will facilitate a more detailed comparison of innovation performance between LA countries and a richer benchmark analysis with economies from the EU. Furthermore, it shed light on the explanatory power of industry and country conditions in the heterogeneity of innovation strategies in the LA context.

Figure 2: Distribution of innovation strategies among firms by size



The innovation practices and strategies here found to have certain common aspects with similar exercise performed in Europe. Specifically, firms pursuing modernization of its production processes relying on the acquisition of external technologies, or those that conduct several innovation practices simultaneously (the high-profile group) represent a considerable share of innovative firms in European and LA countries. On the other hand, the more science-oriented firms commonly found in OECD countries were not noticeable in the LA sample. Furthermore, a considerable group of firms that innovates mainly based on imitation was found, which departs from the innovation strategies reported in similar studies in the EU.

In line with the recent empirical literature, most of the differences in approaches to innovation are due to firm-level heterogeneity. However, the prevalence of imitation strategies is considerable more dependent on country conditions. At the same time, high-profile innovation is not the most common approach followed by the innovative firms in any of the countries here analysed. These aspects raise questions about what is the right STI policy mix to support innovation in the private sector, in the light of the renewed approach to STI policies in LA.

REFERENCES

- Bayona, C., García-Marco, T., & Huerta, E. (2001). Firms' motivations for cooperative R&D: an empirical analysis of Spanish firms. *Research Policy*, 30(8), 1289–1307. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0048733300001517>
- Bell, M., & Pavitt, K. (1992). Accumulating Technological Capability in Developing Countries. *The World Bank Economic Review*, 6(suppl 1), 257–281. Retrieved from http://wber.oxfordjournals.org/lookup/doi/10.1093/wber/6.suppl_1.257
- Cassiman, B., & Veugelers, R. (2006). In Search of Complementarity Internal R & D Strategy: Knowledge in Innovation and External Acquisition. *Management Science*,

- 52(1), 68–82.
- Castellacci, F. (2008). Technological paradigms, regimes and trajectories: Manufacturing and service industries in a new taxonomy of sectoral patterns of innovation. *Research Policy*, 37(6-7), 978–994.
- Crespi, G. A., & Dutrénit, G. (Eds.). (2014). *Science, Technology and Innovation Policies for Development: The Latin American Experience*. Springer International Publishing. Retrieved from <http://link.springer.com/10.1007/978-3-319-04108-7>
- Crespi, G. A., Navarro, J. C., & Zuñiga, P. (2010). *Science, Technology, and Innovation in Latin America and the Caribbean: a Statistical Compendium of Indicators*. Washington DC.
- Crespi, G. A., Tacsir, E., & Vargas, F. (2016). Innovation Dynamics and Productivity: Evidence for Latin America. In I.-A. D. Bank, M. Grazi, & C. Pietrobelli (Eds.), *Firm Innovation and Productivity in Latin America and the Caribbean: The Engine of Economic Development* (pp. 37–71). New York: Palgrave Macmillan US. Retrieved from http://dx.doi.org/10.1057/978-1-349-58151-1_2
- Crespi, G. A., & Vargas, F. (2016). Building an homogenized database of Latin American Innovation Surveys. Manuscript in preparation.
- Crespi, G. A., & Zuniga, P. (2012). Innovation and Productivity: Evidence from Six Latin American Countries. *World Development*, 40(2), 273–290. Elsevier Ltd. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0305750X11001859>
- Frenz, M., & Lambert, R. (2009). Exploring Non-technological and Mixed Modes of Innovation Across Countries. *Innovation in Firms: A Microeconomic Perspective* (OECD., pp. 69–110). Paris: OECD Publishing. Retrieved from http://eprints.bbk.ac.uk/10504/nhttp://www.oecd-ilibrary.org/science-and-technology/innovation-in-firms/exploring-non-technological-and-mixed-modes-of-innovation-across-countries_9789264056213-4-en
- Gault, F. (2010). *Innovation strategies for a global economy: Development, implementation, measurement and management*. International Development Research Centre.
- Hou, J., & Mohnen, P. (2013). *Complementarity between internal knowledge creation and external knowledge sourcing in developing countries* (No. 2013-10). UNU-MERIT Working Paper Series.
- Huang, C., Arundel, A., & Hollanders, H. (2010). *How firms innovate: R&D, non-R&D, and technology adoption* (No. 2010-027). UNU-MERIT Working Paper Series, UNU-MERIT Working Paper Series. Maastricht.
- Jaramillo, H., Lugones, G., & Salazar, M. (2001). *Manual de Bogotá: normalización de indicadores de innovación tecnológica en América Latina y el Caribe*.
- Jensen, M. B., Johnson, B., Lorenz, E., & Lundvall, B. Å. (2007). Forms of knowledge and modes of innovation. *Research Policy*, 36(5), 680–693.
- Leiponen, A., & Drejer, I. (2007). What exactly are technological regimes?. Intra-industry heterogeneity in the organization of innovation activities. *Research Policy*, 36(8), 1221–1238.
- Maloney, W., & Rodríguez-Clare, A. (2007). *Innovation shortfalls. Review of Development Economics* (Vol. 11).
- OECD/Eurostat. (2005). *Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data. OECD and Eurostat Publication* (Vol. Third edit).
- Pavitt, K. (1984). Sectoral patterns of technical change: Towards a taxonomy and a theory. *Research Policy*, 13(6), 343–373.
- Piga, C. A., & Vivarelli, M. (2004). Internal and external R&D: A sample selection approach. *Oxford Bulletin of Economics and Statistics*, 66(4), 457–482.

- Srholec, M. (2011). A multilevel analysis of innovation in developing countries. *Industrial and Corporate Change*, 20(6), 1539–1569.
- Srholec, M., & Verspagen, B. (2012). The Voyage of the Beagle into innovation: Explorations on heterogeneity, selection, and sectors. *Industrial and Corporate Change*, 21(5), 1221–1253.
- Veugelers, R., & Cassiman, B. (2005). R&D cooperation between firms and universities. Some empirical evidence from Belgian manufacturing. *International Journal of Industrial Organization*, 23(5-6), 355–379. Retrieved from <http://linkinghub.elsevier.com/retrieve/pii/S0167718705000251>
- Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal*, 5(2), 171–180. Retrieved from <http://doi.wiley.com/10.1002/smj.4250050207>



Innovation dynamics of Salvadoran agri-food industry from an evolutionary perspective

Elias Humberto Peraza Castaneda^{*}, Guillermo Aleixandre Mendizábal^{**}

^{*} eliashumberto.peraza@alumnos.uva.es

Departamento de Economía Aplicada, Universidad de Valladolid, Av. Valle Esgueva, 6, Valladolid, 47011 (Spain)

^{**} galeixam@eco.uva.es

Departamento de Economía Aplicada, Universidad de Valladolid, Av. Valle Esgueva, 6, Valladolid, 47011 (Spain)

ABSTRACT

This paper presents a holistic approach to analyse the dynamics of innovation of a low-tech sector in a less developed economy, the agri-food industry in El Salvador, in the context of evolutionary economy. This requires using complementary quantitative and qualitative data and methodologies to better understand how Salvadoran agri-food industry innovation system works and how STI public policies can improve the performance of a key sector in terms of national socioeconomic development. The work already done shows a concentrated and vigorous sector with some upstream and downstream connections that innovate depending on firm size, age, R&D activities and use of industrial property rights.

1. INTRODUCTION

Since the definition of national innovation systems was set in late 1980s (Freeman, 1987; Lundvall, 1992; Nelson, 1993), this concept has been widely accepted among researchers dedicated to technological change. Innovation is understood in the terms established by the OECD (2005) and the analysis of innovation systems fits into an evolutionary view of the economy (Nelson y Winter, 1982). From that initial concept, other complementary approaches have been developed to address specific situations related to different territorial dimensions or economic sectors, mainly: regional innovation systems (Cooke *et al.*, 1997) and sectoral innovation systems (Malerba, 2002). The latter, which constitute the framework of this investigation, are "a set of new products and established for specific uses and the set of agents conducting market interactions and no market for the creation, production and sale of these products" (Malerba, 2002, pp. 247). This sectoral approach has been mainly used to study developed countries (Malerba, 2004) although in recent years has also applied to developing countries (Caniëls *et al.*, 2009; Intarakumnerd y Fujita, 2009; Lee, 2009; Malerba y Mani, 2009).

2. INNOVATION IN SALVADORAN AGRIFOOD SECTOR

The agri-food industry is linked both commercially and technologically with the agricultural sector and it is of great importance for the Central American economies, which have historically depended on these two sectors (Soluri, 2009). The Salvadoran economy is not an exception,

with an industrial development marked by the influence of their agricultural export past (Hoselitz, 1954). This sector provided reasonable profit to economic elite during the early 19th century and due to that it did not need to focus on activities with higher added value and innovation till a recent past (Acuña Ortega, 1994). In addition, this sector has been affected by cycles in public policy support and other solid weak support in later decades.

The Salvadoran agri-food industry has an important role in terms of competitiveness, productivity and export capacity. The food industry is the 9.4% of the Salvadoran economy in 2012. Besides, this sector shows a high productivity of labour, 1.27 times the average productivity of the Salvadoran economy for food subsector and, especially, for beverage subsector, with 6.33 times the average. For the second aspect, the agri-food industry has a dynamic export performance, so this sector represented 16.2% of total exports in 2014. Its main commercial destination is Central America and its largest single client is United States.

Nowadays Salvadoran agri-food industry's firms are aware of the need to innovate for keeping their competitiveness in both domestic and international markets. In the case of the agri-food industry innovation refers to:

- Product innovations, such as functional foods (Annunziata y Vecchio, 2011; Jones y Jew, 2007; Sirò *et al.*, 2008)
- Process innovations, as those aimed at promoting the safety, traceability and quality of foods by developing technologies designed to monitor pathogens and other hazards from farm to fork (Aung y Chang, 2014; Caswell *et al.*, 2008; FAO, 2003; U.S. GAO, 2005).
- Organizational innovations, frequently derived from necessary adjustments to meet: ISO quality standards (IOS, 2000; Rao *et al.*, 1997); and requirements of knowledge management systems.
- Marketing innovations, for example: Development of new packaging and product formulations in the face of changing preferences consumers, which is a key element in the food and beverage industry (Tollin, 2008); or the use of origin's designations and other kind of geographical indications (Martínez Ruiz y Jiménez Zarco, 2006).

Bearing in mind the previous context, the objective of the research is to characterise the features of the innovation process of Salvadoran firms in the agri-food industry, a low-tech sector in a less developed economy, within an evolutionary and systemic approach.

3. DATA AND METHODOLOGY

This research uses a multidimensional analysis to integrate the diverse elements that explain the innovative behaviour of firms, which have to be explained from the perspective of the dynamics of learning embedded in the agri-food sectoral systems of innovation, where heterogeneous actors interact in different ways. Therefore, this research is carried out in three complementary points of view:

Firstly, an analysis of the features of Salvadoran economy that influence the innovative behaviour agri-food industry in different ways: a revision of recent history of the agri-food Salvadoran sector allows gathering information to explain the existing situation; the analysis of

sector economic data for the period 2010-2014, and also the analysis of the 2006 input-output table. The economic data comes from the Directorate General of Statistics and the Censuses of the Ministry of Economy and the Central Bank.

Second, a study of innovative behaviour of the agri-food firms to shed light on the determinants of innovations to Salvadoran agri-food firms. This study examines the links between different types of innovation and a set of internal and external factors using multivariate probit regressions (Acosta *et al.*, 2015; Belderbos *et al.*, 2004; Santamaría *et al.*, 2009) utilizing microdata of firm with 10 or more employees. The data comes from the 1st National Innovation Survey 2013 of El Salvador done by the Directorate of Innovation and quality of Economy Ministry.

Thirdly, a survey to relevant public and private actors is needed to know the institutional features that condition the relations of the different player of the sectoral innovation system. So a sectoral system approach examines innovation as the result of both firms' specific variables and the type of knowledge and technologies that characterize a sector, the links and interdependences with other relate sectors, the role of actors such as public agencies and the government, the characteristics of demand and the type of institutions. Also sectoral system approach has a dynamic perspective and it pays a lot of attention to exchange, competition, and cooperation in a co-evolutionary setting. Innovation and diffusion have become relevant in most developing countries (Malerba y Mani, 2009).

4. PRELIMINAY RESULTS AND FUTURE CHALLENGES

The work done so far shows some preliminary results: The Salvadoran agri-food industry shows a territorial concentration in two departments: San Salvador and La Libertad, followed at a great distance by the department of Sonsonate, adjacent to La Libertad. Besides, these three departments and Usulután reveal a specialisation profile in agri-food industry activities. Based on the foregoing, possible support actions can focus on these territories, although authorities must bear in mind to address the asymmetries of access to opportunities for actors located outside these agglomerations, which constrain the social inclusion of the innovation system (Dutrénit y Sutz, 2014).

The agri-food industry has some upstream and downstream commercial connections with other sectors: backward relations link with the industry itself, along with the livestock and basic grains sector; and forward relations link with the poultry sector, the livestock, the restaurant and hotel sector and also the agri-food industry itself. Therefore, competitive related measures need to be coordinated with agricultural policies holistically; to take into account other chained sectors. Besides, technological similarities are identified between agri-food industry and poultry sector.

The result of the multivariate probit regression based on the data of 378 agri-food firms shed light on the internal and external determinants of innovation, which depends on the type of innovation. Product innovation is positively related with R&D activities, use of industry property right, firm's size by sales, internal information sources, knowledge agents' information sources, urbanization economies, low public policy incentives; and negatively

related with age, and exports. Process innovation is positively related with age, internal information sources, knowledge agents' information sources, value chain agents' information sources; and negatively related with firm's size by sales. Organisational innovations are positively related with university employees' percentage, internal information sources, knowledge agents' information sources, value chain agents' information sources; and negatively related with exports, urbanization economies and low public policy incentives. Finally, marketing innovations are positively related with firm's size by sales, university employees' percentage, R&D activities, internal information sources, knowledge agents' information sources; and negatively related with exports.

The previous results provide empirical evidence on the performance of agri-food sector and on the determinants of the innovative behaviour of their firms for the case of a developing economy. This can be useful for Salvadoran Public Administration in the design agri-food innovation policies, e.g. an agri-food cluster initiative.

The future challenge is to integrate this quantitative data with two complementary sources: on the one hand, qualitative data coming from the study of the recent political and economic history of agri-food industry and the institutional framework where firms operate and, also, from a survey to relevant public and private actors of the agri-food innovation system; on the other hand, the main economy and technological trends that characterised the innovation of agri-food sector in the global markets that constrain the activity agri-food Salvadoran sector.

REFERENCES

- Acosta, M.; Coronado, D.; y Romero, C. (2015). Linking public support, R&D, innovation and productivity: New evidence from the Spanish food industry. *Food Policy*, 57, 50-61. doi: <http://dx.doi.org/10.1016/j.foodpol.2015.09.005>
- Acuña Ortega, V. H. (1994). *Historia General de Centroamérica. Tomo IV: Las Repúblicas Agroexportadoras (1870 - 1945)*. San José, Costa Rica: Facultad Latinoamericana de Ciencias Sociales. FLACSO.
- Annunziata, A.; y Vecchio, R. (2011). Functional foods development in the European market: a consumer perspective. *Journal of Functional Foods*, 3, 223-228. doi: <http://dx.doi:10.1016/j.jff.2011.03.011>
- Aung, M. M.; y Chang, Y. S. (2014). Traceability in a food supply chain: Safety and quality perspectives. *Food Control*, 39, 172-184. doi: <http://dx.doi:10.1016/j.foodcont.2013.11.007>
- Belderbos, R.; Carree, M.; Diederer, B.; Lokshin, B.; y Veugelers, R. (2004). Heterogeneity in R&D cooperation strategies. *International Journal of Industrial Organization*, 22, 1237-1263. doi: <http://dx.doi.org/10.1016/j.ijindorg.2004.08.001>
- Caniëls, M.; Kesidou, E.; y Romijn, H. (2009). The Software Sector in Uruguay: a sectoral systems of innovation perspective. In F. Malerba y S. Mani (Eds.), *Sectoral Systems of Innovation and Production in Developing Countries. Actors, Structure and Evolution*. Cheltenham: Edward Elgar.
- Caswell, J.; Roberts, T.; Golan, E.; y Salay, E. (2008). The Interaction of Public and Private Incentives in Promoting Food Safety Innovation in the U.S. Meat Industry. In R. Rama

- (Ed.), *Handbook of innovation in the food and drink industry*. New York, United States of America: The Haworth Press. Taylor & Francis Group.
- Cooke, P.; Gómez, M.; y Etxebarria, G. (1997). Regional Innovation Systems: Institutional and Organizational Dimensions. *Research Policy*, 26(4-5), 475-491. doi: [http://dx.doi:10.1016/S0048-7333\(97\)00025-5](http://dx.doi:10.1016/S0048-7333(97)00025-5)
- Dutrénit, G.; y Sutz, J. (2014). *National Innovation Systems, Social Inclusion And Development. The Latin American Experience*. Northampton, Massachusetts: Edward Elgar.
- FAO. (2003). *FAO's strategy for a food chain approach to food safety and quality: A framework document for the development of future strategic direction*. Rome: Food and Agriculture Organization.
- Freeman, C. (1987). *Technology Policy and Economic Performance: Lessons from Japan*: London: Pinter.
- Hoselitz, B. F. (1954). *Industrial Development of El Salvador*. Washington D.C.: Catholic University of America.
- Intarakumnerd, P.; y Fujita, M. (2009). China's threat and opportunity for the Thai and Vietnamese motorcycle industries: a sectoral innovation system analysis. In F. y. M. Malerba, S. (Ed.), *Sectoral Systems of Innovation and Production in Developing Countries. Actors, Structure and Evolution*. Cheltenham: Edward Elgar.
- IOS. (2000). *ISO 9001*. Geneva: International Organization for Standardization.
- Jones, P. J.; y Jew, S. (2007). Functional food development: concept to reality. *Trends in Food Science & Technology*, 18, 387-390. doi: <http://dx.doi:10.1016/j.tifs.2007.03.008>
- Lee, T.-L. (2009). From nuts and bolts to bits and bytes: the evolution of Taiwan ICT in a global knowledge-based economy. In F. Malerba y S. Mani (Eds.), *Sectoral Systems of Innovation and Production in Developing Countries. Actors, Structure and Evolution*. Cheltenham: Edward Elgar.
- Lundvall, B. Å. (1992). *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning*: London: Pinter.
- Malerba, F. (2002). Sectoral Systems of Innovation and Production. *Research Policy*, 31, 247-264. doi: [http://dx.doi.org/10.1016/S0048-7333\(01\)00139-1](http://dx.doi.org/10.1016/S0048-7333(01)00139-1)
- Malerba, F. (2004). *Sectoral Systems of Innovation: Concepts, Issues and analyses of Six Major Sectors in Europe*. Cambridge: Cambridge University Press.
- Malerba, F.; y Mani, S. (2009). *Sectoral Systems of Innovation and Production in Developing Countries*. Cheltenham: Edward Elgar.
- Martínez Ruiz, M. P.; y Jiménez Zarco, A. I. (2006). La potenciación del origen en las estrategias de marketing de productos agroalimentarios. Objetivos, situación e implicaciones. *Boletín ICE Económico*, N°2880, 13-29.
- Nelson, R. R. (1993). *National Systems of Innovation: A Comparative Study*. Oxford: Oxford University Press.
- Nelson, R. R.; y Winter, S. G. (1982). *An Evolutionary Theory of Economic change*. Cambridge, Massachusetts: Harvard University Press.
- OCDE. (2005). *Manual de Oslo: Guía para la recogida e interpretación de datos sobre innovación*. París: OCDE.

- Rao, S. S.; Ragu-Nathan, T. S.; y Solis, L. E. (1997). Does ISO 9000 have an effect on quality management practices? An international empirical study. *Total Quality Management*, 8, 335-346. doi: <http://dx.doi:10.1080/0954412979352>
- Santamaría, L.; Nieto, M. J.; y Barge-Gil, A. (2009). Beyond formal R&D: Taking advantage of other sources of innovation in low- and medium-technology industries. *Research Policy*, 38, 507-517. doi: <http://dx.doi.org/10.1016/j.respol.2008.10.004>
- Sirò, I.; Kàpolna, E.; Kàpolna, B.; y Lugasi, A. (2008). Functional food. Product development, marketing, and consumer acceptance - a review. *Appetite*, 51, 456-467. doi: <http://dx.doi:10.1016/j.appet.2008.05.060>
- Soluri, J. (2009). *Banana Cultures: Agriculture, Consumption, and Environmental Change in Honduras and the United States*. Austin, Texas, United States of America.: University of Texas Press.
- Tollin, K. (2008). Managing Knowledge for Product Innovation. In R. Rama (Ed.), *Handbook of innovation in the food and drink industry*. New York, United States of America: The Haworth Press. Taylor & Francis Group.
- U.S. GAO. (2005). *Overseeing the U.S. food supply: Steps should be taken to reduce overlapping inspections and related activities*. . Washington, D.C.: U.S. Government Accountability Office.



Elucidate Innovation Performance of Technology-driven Mergers and Acquisitions

Lu Huang*, Kangrui Wang*, Huizhu Yu*, Lining Shang*, Liliana Mitkova**

*huanglu628@163.com;kangruiwbb@163.com;675244410@qq.com; sln_work@163.com; School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China

**Liliana.Mitkova@u-pem.fr; Univ Paris Est Marne Vallee, F-77420 Champs Sur Marne, France

ABSTRACT

The importance and value of Mergers and Acquisitions (M&As) have increased with the expectancy to obtain key technology capabilities and rapid impact on innovation. This article develops an original analytical framework to elucidate the impact of the technology and product relatedness (similarity/complementarity) of the Technology-driven M&A' partners on post-innovation performance. We present results drawing on a multiple case studies of Chinese High-Tech firms from three industries.

Keywords

Innovation Performance, Technology-driven Mergers and Acquisitions, Technology and product relatedness

INTRODUCTION

Recently, many world-class companies, especially those in high technology industries, have leaped forward through technology-driven mergers and acquisitions (Tech M&As). For example, Apple acquired fingerprint sensor maker AuthenTec, and launched the iPhone5S with fingerprint recognition in 2013 successfully.

The key role of Tech M&As was underlined as an important way to booster the companies' competitiveness through the acquisition of an external knowledge resources (Ferreira& Santos, 2014). It is important to understand the impact of M&As on the post-innovation performance as a complex result revealing how firms absorb and use external knowledge, how they produce innovation outputs from market point of view, and how they redeploy resources form resources-based view (Ahuja and Katila, 2001). However, present studies still stay largely limited from different points of view: 1) exclusively focusing on one particular outcome, such as, R&D input & output, R&D efficiency, or new product development (Valentini, 2012); 2) concluding with general positive/negative judgements (Pervan and al. 2015), and ignoring to analyse the contribution of essential parameters to achievement of effective innovation performance for Tech M&As; 3) focusing on Tech M&As' impact in USA and European companies, neglecting the new insights from emerging context.

The study develops an original analytical framework to elucidate the impact of the technology and product relatedness (similarity/complementarity) of the Tech M&A' partners on post-innovation performance. The performance is appreciated via complex parameters involving R&D input, patent & product activity, and financial results from commercialization. To illustrate the implementation of this method, the empirical study draws on a multiple case study of three different industries in China including traditional manufacturing, emerging technology and an interdisciplinary high tech sector.

METHODS

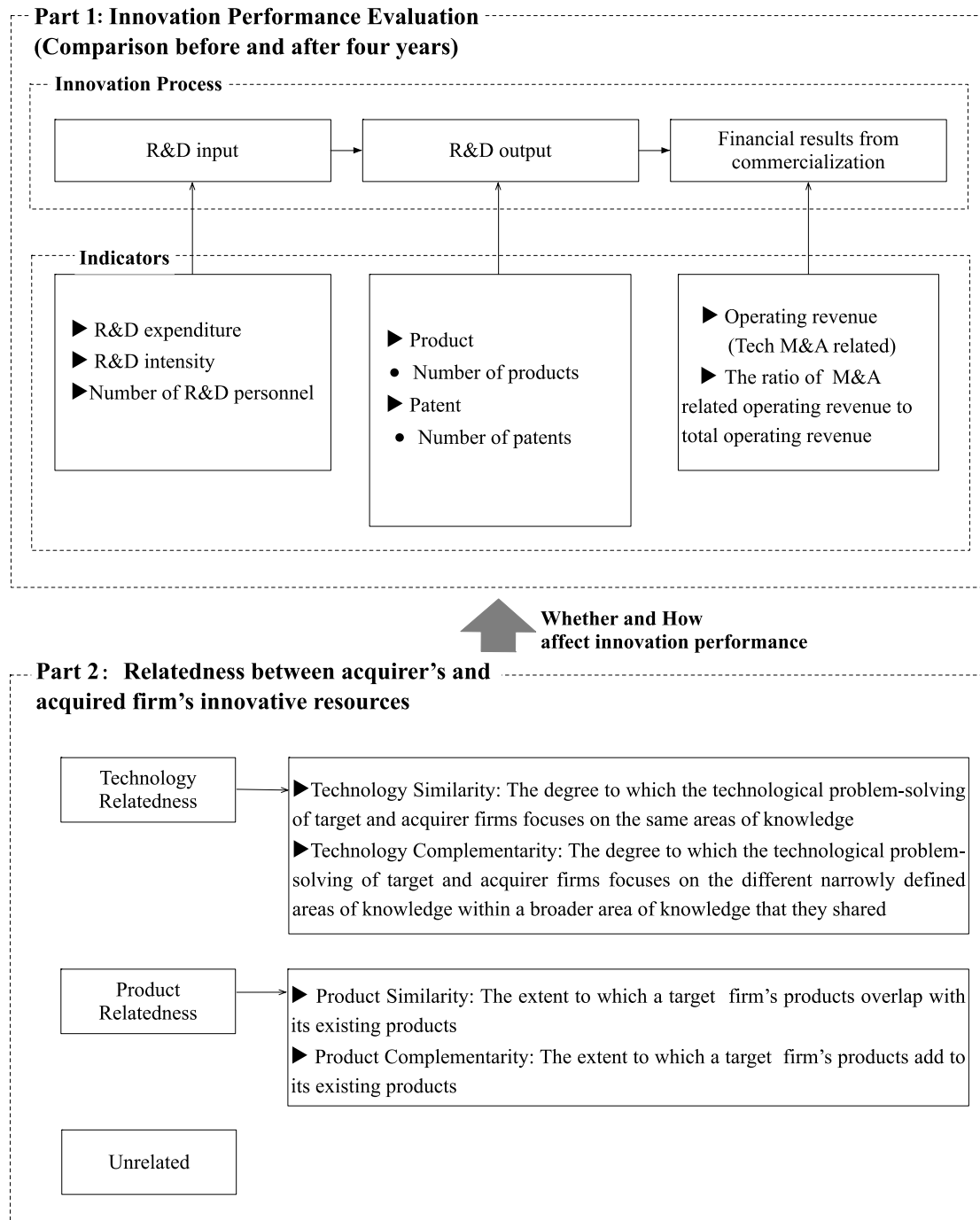
Existing researches propose variety of studies addressing Tech M&As innovation performance in terms of operational indicators, variables, temporal orientation, methods, and data. However, these contributions are fragmented using one-dimensional indicators to analyse innovation performance, taking advantage of isolated perspective to explore the relatedness between M&A partner. In order to contribute to the building of a comprehensive method to assess innovation performance of Tech M&As, the objective of this study is to suggest a new framework guided by the following driving questions:

1. How to explore the similarity and complementarity between acquirers and acquired firms in terms of technology and product relatedness?
2. How to create multi-dimensional indexes taking into account R&D activity and financial results from commercialization to evaluate and compare target/acquiring firms' innovation performance before and after Tech M&A transaction?

We devise a new methodological framework to elucidate the link between technology and product relatedness of partners and M&As post-innovation performance. The proposed framework consists of two parts (as shown in Figure 1).

The first part (part1) focuses on evaluating the post-innovation performance from the entire innovation process perspectives. It's important to indicate that all the indicators in Part 1 are referred to the Tech M&A-related business sector, not including the whole sectors of the acquirer firm. In order to reveal the impact of M&As on these complex parameters, the comparison analysis would be conducted between the innovation performance before and after four years of the time of M&A transaction occurs.

Figure 1: Tech M&As Innovation Performance Evaluation Design



The second part (part 2) of the proposed framework is dedicated on the investigation of the relatedness type of M&As partners' innovative resources, and the exploration of the link between these relatedness and innovation performance.

According to resources-based view (Barney 1991, Teece et al. 1997), two significant innovation-based resources represent attractive opportunities for M&As: technological incomes (innovation inputs) and commercialized products (Innovation output) (Yu et al. 2015). Many researches clarify that M&A in related fields would have better performance than in non-related fields (Barney 1991; John and Harrison 1999). At the same time, the outcome of the innovation process is the product which is a source of revenues and instrument to penetrate an existing or new market (Lichtenthaler and Ernst 2012). Therefore, both technology relatedness and product relatedness of the partners may affect the innovation process and innovation performance. Consequently, in the proposed framework we consider technology and product relatedness by the similarity or complementarity.

Technology similarity

We calculate the technology similarity via an IPC-based categorical similarity measure approach (Zhang et al., 2016). The core concept of technology similarity measurement is to denote IPC code of each firm as a fuzzy set, and then use Cosine function to measure the categorical similarity between the two patent portfolios of the two firms. Technology similarity between Firm a and Firm b could be measured below:

$$TS(a, b) = \frac{V(a) \times V(b)}{|V(a)||V(b)|}$$

Where $V(a)$ is a m -dimensional vector $V(a) = \{\vartheta_1, \vartheta_2, \dots, \vartheta_j, \dots, \vartheta_{m-1}, \vartheta_m\}$, and ϑ_j is the membership grade that the firm a belongs to the fuzzy set A_j . The membership function $A_j(a)$ is considered as the degree with which firm a engages IPC A_j .

Technology complementarity

We follow Makri et al. (2010) to construct measures of technology complementarity based on patent data of the acquirer and target firm. Technology complementarity is calculated using the number of patents in the same category but in different patent classes, and given by:

$$TC(a, b) = \frac{\text{Overlap all patent categories}}{\text{Total patents a\&b}} - \left(\frac{\text{Overlap all patent class}}{\text{Total patents a\&b}} \times \frac{\text{Total acquirer patents in common categories}}{\text{Total acquirer patents}} \right)$$

The measures of technology complementarity are weighted by the importance of each patent class for the acquirer, in order to account for the fact that large firms tend to patent in various patent classes (Miozzo and DiVito et al., 2015).

We apply Yu's theory and measurements for the classification of M&A from the product perspective (Yu et al., 2015).

Product Similarity:

$S_{p,p}(i, j)$, for acquirer i and target firm j measures the similarity of the two firms' products using the following expression:

$$S_{p,p}(i, j) = \sum_i \sum_j w(C_o, C_m) [a \cdot s_p(C_o, i) + b \cdot s_p(C_m, j)]$$

Here $s_p(C_o, i)$ is the product score, means the market potential that firm i has in each area C_o .

a and b are constants that satisfy the constraints: $0 < a, b < 1$, and $a + b = 1$. The weight function, $w(C_o, C_m)$, reflects the distance between two areas.

Product complementarity:

$E_{p,p}(i, j)$, is defined as:

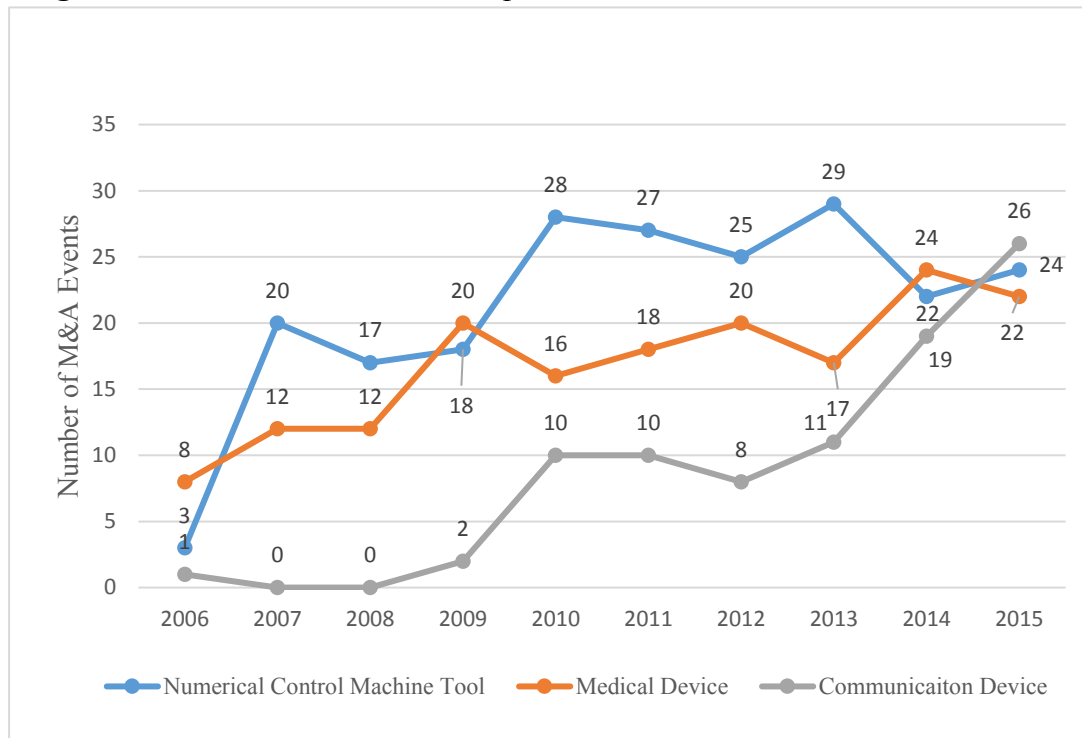
$$E_{p,p}(i, j) = \sum_j r(C_p^i, C_m) s_p(C_m, j)$$

Here $r(C_p^i, C_m)$ is a weight function influenced by the distance between C_m and any area that the acquirer's products are belonged to.

At the end, we compare the indicators of all acquirers four years before and after the transactions with the significance of Paired-Samples t-Test (Raju, 2005) to check if the difference is significant under some confidence interval.

DATA

We selected three representative Chinese industries: computer numerical control machine tool (CNCMT), medical device and communication device. This selection was based on three main reasons. Firstly, "Made in China 2025" plan clearly put forward ten key high technology areas which would be strongly reinforced in the next ten years. In these areas, new information technology, high-end numerically-controlled machine tools and robotics, and high performance medical device are listed. Secondly, the analysis of the numbers of M&A transactions in these three industries from 2006 to 2015 (as shown in Figure 2) illustrates that the three industries are active in M&A events.

Figure 2: M&A Transactions Development Trend of the three industries from 2006 to 2015

Thirdly, the selected industries are “high technology” but with different development trend. Numerical control machine tool is as traditional manufacturing; communication device as an emerging technology industry and medical device industry as interdisciplinary high tech sector. This diversity allows getting valuable insights from the comparison of innovation performance results.

The Tech M&A database was built for the selected three industries in China. According to Ahuja (2001), the selection rule is whether the technology is a motivating factor for the acquisition or if the technology is a part of the transferred assets mentioned in the acquisition announcements. In order to eliminate other factors’ effect and focus on the innovation performance deducted by one Tech M&A transaction, we set a nine-year period as the investigation time and there is only one Tech M&A event occurs in the middle year (the fifth year) and none in the other eight years. Thus, a comparison of the innovation performance before and after four years of the time of Tech M&A transaction occurred could be made, which allows accurate judgements about the effects of Tech M&A (Man & Duysters, 2005). Furthermore, we tried to select the M&A transactions occurring in the same year for each industry, in order to ensure the same industry environment. According to the above selection criteria, 15 Tech M&A cases were selected finally, including 5 cases in NCMT industry occurring in the year of 2010, and 5 cases in both medical device and communication device industry occurring in the year of 2011 (Table 1,2,3).

Table 1. Presentation of the 5 Tech M&A cases in Computer Numerical Control Machine industry

Industry	CASE	Acquirer	Acquired Firm	M&A Amount	M&A Purpose
Computer Numerical Control Machine Tool (CNCMT)	1	Chongqing Machinery & Electric Co., Ltd. (CQME)	Precision Technologies Group Ltd. (PTG)	\$28,552,000	CQME recognized the strengths and advantages inherent in PTG's market knowledge and technologies, which over the years has enabled PTG to become world-leaders. Through M&A, CQME saw the opportunity to add value and complementary products to its already extensive product range, the highly developed technologies of Holroyd Precision to enter new markets in the global precision gear manufacturing sector.
	2	MESNAC Co., Ltd. (MESNAC)	Qingdao Kingerbot Automation Co., Ltd. (KINGER)	\$1,108,634	MESNAC aimed to advanced automation technologies of KINGER. Through M&A, MESNAC could increase its automation level and improve the quality of its products.
	3	SIASUN Automation Co., Ltd (SIASUN)	SHENHUI TT&C Co., Ltd. (SHENHUI)	\$923,400	SHENHUI is famous for its advanced servo motor technologies of industrial robot. Through M&A, SIASUN would like to realize the self-production of robot accessories, which is a new emerging business sector of SIASUN.
	4	JINSHENG Group (JINSHENG)	EMAG Holding GmbH (EMAG)	\$111,270,000	EMAG is a global leading CNCMT company with advanced technologies. Through M&A, JINSHENG could enter CNCMT industry.
	5	Zhejiang TianMa Bearing Co., Ltd. (TIANMA)	Qiqihar Heavy CNC Equipment Co., Ltd. (QIQIHAR)	\$18,846,748	QIQIHAR is a high-tech company in Bearing industry, which is the upstream of Machine Tool industry. Through M&A, TIANMA could further extend the bearing manufacturing industry chain, and strengthen the link between upstream and downstream industries.

This work is licensed under a Creative Commons License: Attribution-NonCommercial-NoDerivatives 4.0 International.

Table 2. Presentation of the 5 Tech M&A cases in Medical Device industry

Industry	CASE	Acquirer	Acquired Firm	M&A Amount	M&A Purpose (from M&A announcement document)
Medical Device	1	Zhejiang Hongda Warp Knitting Co., Ltd. (HONGDA)	Shenzhen Well.D Medical Electronics Co., Ltd. (WELLD)	\$55,323,972	HONGDA is in textile industry and this acquisition will help it enter the medical device industry.
	2	Inner Mongolia Furui Medical Science Co., Ltd (MONGOLIA)	Echosens S.A.S. (ES)	\$22,254,000	MONGOLIA is a chemical pharmaceutical enterprise for the treatment of liver diseases. ES company mainly produces liver elasticity testing equipment "Fibroscan". Through M&A, MONGOLIA could realize the system of "diagnosis plus treatment" of liver disease.
	3	Guangdong Biolight Meditech Co., Ltd. (Biolight)	Tianjin ever-trust medical equipment development Co., Ltd. (EVER)	\$2,924,100	Through M&A, Biolight would like to extend its medical monitoring equipment product line, obtaining the advanced technologies in the field of hemodialysis from EVER.
	4	MicroPort Scientific Corporation (MicroPort)	Dongguan Kewei Medical Instrument Co., Ltd. (Kewei)	\$16,621,200	Both MicroPort and Kewei are in medical device industry. MicroPort could plug the gap in its product line of cardiac surgery by the acquisition.
	5	Edan Instruments, Inc. (Edan)	Shenzhen Unitelab Bio-electric Co. Ltd (Unitelab)	\$2,154,600	Edan and Unitelab are in the same field of clinical analytical instruments. Through M&A, Edan could obtain advanced blood cell analysis instrument from Unitelab, and upgrade Edan's technical level.

Table 3. Presentation of the 5 Tech M&A cases in Communication Device industry

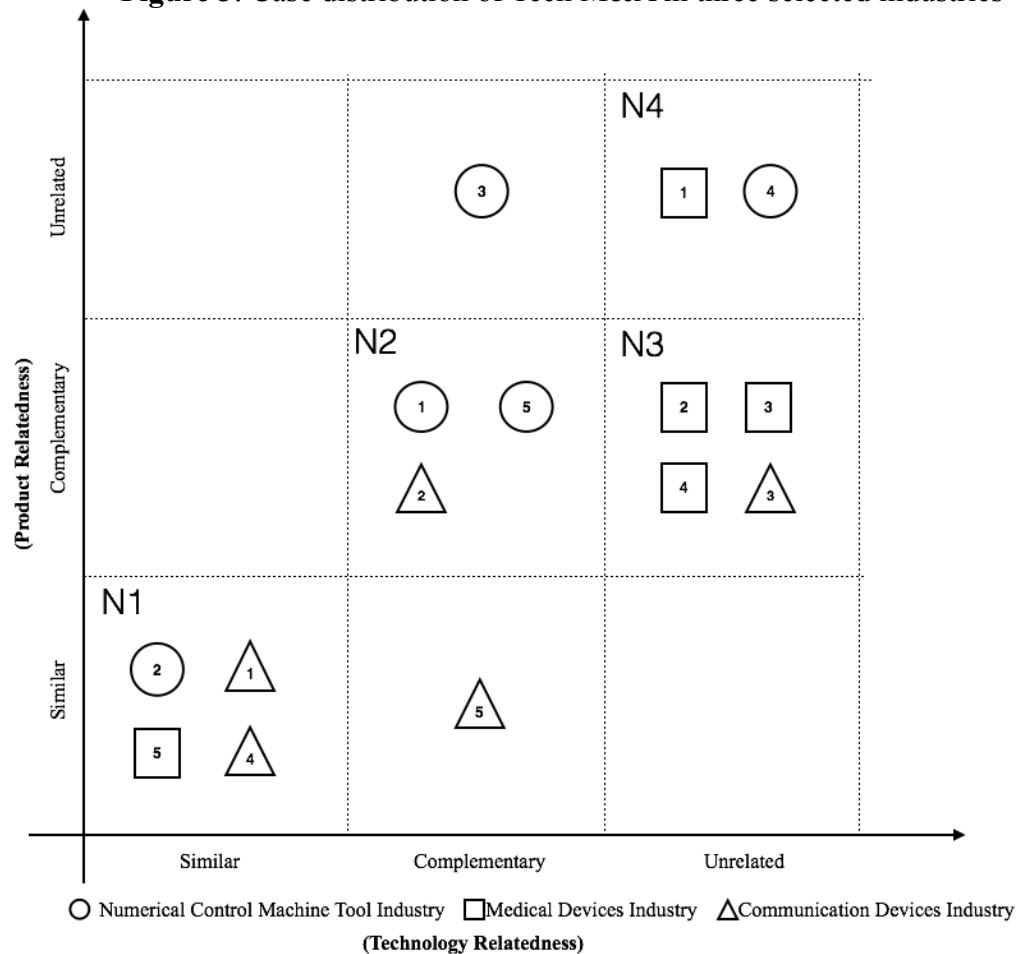
Industry	CASE	Acquirer	Acquired Firm	M&A Amount	M&A Purpose (from M&A announcement document)
Communication Device	1	SUNWAVE Solutions (SUNWAVE)	Guangzhou Eshine Electronic Technology Co., Ltd (Eshine)	\$5,484,996	SUNWAVE and Eshine are in the same field of mobile communication network optimization. Through M&A, SUNWAVE could greatly enhance its network optimization capability.
	2	Hytera Communications Co. Ltd (Hytera)	Rohde & Schwarz Professional Mobile Radio GmbH ((PMR))	\$2,225,400	PMR's TETRA infrastructure solution is a good compliment to Hytera's TETRA terminal product portfolio. The acquisition enables Hytera to offer complete TETRA system solutions to its customers.
	3	Hangzhou Huaxing Chuangye C.T.S Co.,Ltd (Huaxing)	Shanghai Xinzhong C.T Co., Ltd (XZCOM)	\$6,639,246	Through M&A, Huaxing extended to the end of the mobile network optimization industry chain, and began to get involved in network equipment and optimization solutions business.
	4	Dingli Shiji Co.,Ltd (Dingli)	Bettercomm Co.,Ltd (BCM)	\$3,808,256	Through M&A, Dingli could improve the ability of network optimization services, and strengthen its technology power.
	5	North Navigation Control Technology Co., Ltd (NNCT)	Henan Wanxiang C.T Co., Ltd (WXCT)	\$13,778,221	Through M&A, NNCT could obtain the advanced VHF/ UHF aviation communication technologies and digital satellite mobile communication technologies from WXCT, greatly enhancing its technical capability in the field of military communications.

RESULTS*Measuring technology and product relatedness of the Tech M&A cases*

We collected the patent information of 30 companies (the acquirer and the acquired firms of the 15 M&A cases) to calculate their technology similarity and technology complementarity. Meanwhile, we also constructed the product system for each company in order to measure the product similarity and product complementarity as indicated in the third methodological part. The analyzed results of the technology relatedness and product relatedness for the acquirers and the targets in the selected cases are illustrated in Table 4.

Table 4. Technology and Product Relatedness of the Case Studies in the three industries

Industry	CASE	Technology similarity	Technology complementarity	Technology Relatedness	Product similarity	Product complementarity	Product Relatedness
Numerical Control Machine Tool (NCMT)	1	0.137	0.429	Technology complementary	0.167	0.325	Product complementary
	2	0.613	0.05	Technology similar	0.357	0.004	Product similar
	3	0.032	0.378	Technology complementary	0	0	Product unrelated
	4	0	0	Technology unrelated	0	0	Product unrelated
	5	0.047	0.442	Technology complementary	0.17	0.455	Product complementary
Medical Device	1	0	0	Technology unrelated	0	0	Product unrelated
	2	0.013	0.002	Technology unrelated	0.036	0.667	Product complementary
	3	0.064	0.007	Technology unrelated	0.007	0.378	Product complementary
	4	0.028	0.016	Technology unrelated	0.102	0.326	Product complementary
	5	0.707	0	Technology similar	0.314	0.002	Product similar
Communication Device	1	0.411	0	Technology similar	0.459	0.021	Product similar
	2	0.013	0.422	Technology complementary	0.017	0.435	Product complementary
	3	0.09	0.002	Technology unrelated	0.013	0.372	Product complementary
	4	0.408	0.104	Technology similar	0.384	0.012	Product similar
	5	0.002	0.382	Technology complementary	0.438	0.164	Product similar

Figure 3: Case distribution of Tech M&A in three selected industries

The matrix in Figure 3 represents a synthesis of the results of table 4. In the matrix, the number in the symbol represents the total case number of specific industry in the given relatedness type.

The results illustrate that the majority of M&A relatedness types includes:

1) Product similar & Technologically similar partners (Square N1): the purpose of this type of M&A is to enhance the core competitiveness within acquirer's present technology domain. The key point is to become stronger via acquiring prominent high-tech firms in the given industry of the acquirer. In this situation, the M&A partners share similar technology and product base;

2) Product complementary & Technologically complementary/unrelated partners (Square N2 and Square N3): the purpose of this type of M&A is to expand the current business scope and scale, and rich product range. For example, big companies tend to acquire targets in different positions of the industrial chain to complement their product lines. The technology bases of the acquirer and the acquired target for such kind of M&As maybe complement each other or be unrelated;

3) Product unrelated & Technology unrelated partners (Square N4): the purpose of this type of M&A is to enter a new field, and the acquirer could obtain advanced technologies which are unrelated with its current technology system.

Exploring innovation performance of the Tech M&A cases

In this sector, we aim to explore the innovation performance of Technology-driven M&A cases according to the indicators which are set in Fig. 2, and sum up post-M&A performance conclusions based on different technology and product relatedness types. Fig. 3 shows that there are 13 cases distributed in the four types, and we mainly chose these major types (Type N1, N2, N3, and N4) to compare for their representativeness. Paired sample t-test was utilized to examine whether there were statistically significant differences between the indicators four years before and after the transactions, in order to check whether Technology-driven M&A transaction plays a significant role in the innovation performance of the acquirer.

*R&D input***Table 5.** t-value of R&D input indicators for four M&A relatedness types

Variable	Product similar& Technologically similar (N1) ^a	Product complementary& Technologically complementary (N2) ^a	Product complementary& Technologically unrelated (N3) ^a	Product unrelated & Technologically unrelated (N4) ^a
R&D expenditure	-8.757 ^{**}	-1.417	-8.184 ^{**}	-11.853 ^{**}
R&D intensity	-2.691 [*]	-0.318	-0.600	-0.443
Number of R&D personnel	-3.781 ^{**}	0.065	-2.466	-18.426 ^{**}

Notes: ^a Mean values; standard errors in parentheses.

^{*} P<0.05; ^{**} P<0.01. These are two-tailed significance levels using robust standard errors.

The results in Table 5 indicate that the R&D input of the acquirer increases after Technology-driven M&A as a whole. Only Type N2 seems have a negative performance on the R&D input, especially on the R&D personnel input. This is because the M&A targets of this type are mostly mature technologies or products, and the acquirer tends to take advantage of M&A to use them directly other than conduct new R&D activities.

In comparison, Type N1 performances best on the R&D input with significant confidence level in all three indicators. We further examine the raw data and found that all four cases of this type have achieved a double value of R&D intensity in the fourth year after M&A. Technology or product in the similar area would facilitate the acquirer easily digest the acquired technologies, and quickly start new and mega R&D projects.

Considering R&D intensity, this factor in medical device industry is higher than other two industries. The average value of the five cases in the medical device is more than 8%, however, the one in the NCMT industry is between 2% and 4%, while the one in the communication device is between 6% and 10%, which shows that emerging industry tends to increase the R&D investment after Technology-driven M&A.

R&D output

In this step, patent and product activity is the second parameter of M&A innovation

performance. We collected the raw data of patent number and patent citation from Derwent database and gathered product data from annual reports. The t-test results of R&D output and application indicators for four M&A relatedness types is shown in Table 6.

Table 6 T-value of R&D output and application indicators for Four M&A relatedness types

Variable	Product similar & Technologically similar (N1) ^a	Product complementary & Technologically complementary (N2) ^a	Product complementary & Technologically unrelated (N3) ^a	Product unrelated & Technologically unrelated (N4) ^a
Number of patent	-1.145 ^{**}	-1.177	1.075	0.970
Number of product	-4.219 ^{**}	0.402	-2.000	-1.528
Number of citation	-2.235	-3.261	2.417	-4.333

Notes: ^a Mean values; standard errors in parentheses.

^{*} P<0.05; ^{**} P<0.01. These are two-tailed significance levels using robust standard errors.

Table 6 indicates that the effect of M&A on the R&D output does not have a real tendency since there are both some negative numbers and some positive numbers. In general, the existing studies (Cloodt *et al.*, 2006; Makri *et al.*, 2010) show that M&A events could have either positive or negative impacts on invention performance. Furthermore, compared with table 5, the absolute value of the minus number in table 8 is much lower, which means that the R&D output increment is not as significant as the one of R&D input. That is because the achievement of R&D output requires a certain period of time, and the effect of M&A only could show up in the next few years.

The results in table 8 also reveal that type N1 still have better performances on the R&D output compared with other three types. Meanwhile, type N3 has a poor behavior in the patent outcome (scores are positive numbers) but perform not bad in the new product, because it's difficult to digest unrelated technologies and come up with new patents, but complementary product may help enhance acquirer's product line.

*Financial results from commercialization***Table 7.** T-value of Financial results for Four M&A relatedness types

Variable	Product similar & Technologically similar (N1) ^a	Product complementary & Technologically complementary (N2) ^a	Product complementary & Technologically unrelated (N3) ^a	Product unrelated & Technologically unrelated (N4) ^a
Technology-driven M&A related Operating Revenue	-1.632	-1.142	-4.538 ^{**}	-5.367 [*]
The ratio of M&A related operating revenue to total operating income	-1.312	-1.039	-2.585 [*]	-6.328 ^{**}

Notes: ^a Mean values; standard errors in parentheses.

^{*} P<0.05; ^{**} P<0.01. These are two-tailed significance levels using robust standard errors.

Table 7 shows that the financial Results from Commercialization of the acquirer increases after Technology-driven M&A as a whole, especially, Type N3 and Type N4 have better performances on commercialization than type N1 and type N2.

Most cases of Type N3 and Type N4 are in medical device industry, whose aim of these Technology-driven M&A cases is to either enter this industry for a freshman or enter new sub-sectors for the player already in this industry, and the new business acquired could generate profits immediately. However, similar technology and similar product M&A would focus on developing new R&D project and new products, which need more time for realizing commercialization profits; while for technologically complementary and product complementary M&A, the improvement of profitability through technology integration also needs more time.

Overall Innovation Performance Evaluation

Based on above respective analysis and their variance contributions, we conclude the innovation performance for the major four M&A relatedness types into three levels in table 8.

Table 8. Innovation Performance of Different Types

TYPE	Technology and Product Relatedness	R&D input	R&D output and application	Financial results from commercialization
N1	Product similar & Technologically similar	high	high	medium
N2	Product complementary & Technologically complementary	low	low	medium
N3	Product complementary & Technologically unrelated	medium	low	high
N4	Product unrelated & Technologically unrelated	high	low	high

Notes: High means the p values of all the indicators are significant;

Medium means that the p values of more than one indicators are significant, and all the t-values are minus;

Low means that the t-values of more than one indicators are positive.

It is apparent that Type N1 could lead to a good performance both in R&D input and in R&D output and application. Type N3 and Type N4, which are both independent in technology and aiming to enter a new business area or a new sub-sector, can have better results in commercialization profits; in addition, unrelated product & unrelated technology M&A (Type N3) also could bring high R&D input. In comparison, Type N2 is always reflected as the vertical M&A, the acquirer company of which always acquires the target in different positions of the industrial chain to complement the product lines, does not have significant effect on the innovation performance.

CONCLUSION

This study develops an analytical framework that elucidates the impact of the technology and product relatedness of the Tech M&As' partners on post-innovation performance from three perspectives: R&D input, patent & product activity, and financial results from commercialization.

There are three contributions in this study. Firstly, a multi-dimensional framework is proposed, by taking advantage of a quantitative methodology allowing to analyse the link between partners' technology and product relatedness and Technology-driven M&As innovation performance. Secondly, our approach compares the innovation performance before and after Technology-driven M&A, taking into account the whole innovation process from R&D input to commercialization output. Thirdly, empirical analysis in three different industries (traditional manufacturing, emerging technologies and an interdisciplinary high-tech sector) of the emerging market – China- also provides new insights from a “catching-up” development strategy country.

Several limitations of this study are worth noting. Because of the strict rules of selecting Technology-driven M&A transactions, our study only examined 15 cases in the three industries. The restriction of the low number of observations reinforces the need for conducting the study in more other industries. Furthermore, we neglect the

effect of firm size on the patent and product activities, and the industrial concentration is also an important factor for the decision of M&A, which should be considered comprehensively in the future study.

This study has tried to figure out the links between the relatedness of Technology-driven M&A partners and the post-M&A innovation performance, however, the questions of “why” do these links happen have not be explored. In our next work, we would track R&D-related firm behaviors of the acquirer after M&A and conclude the insights on what kind of behaviors would have better performance for the given M&A relatedness type.

REFERENCES

- Ahuja, G., and Katila, R. (2001) “Technological acquisitions and the innovation performance of acquiring firms: A longitudinal study.” *Strategic Management Journal*, Vol. 22, pp. 197–220.
- Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of management*, 17(1), 99-120.
- Cloodt M, Hagedoorn J, Van Kranenburg H. 2006. Mergers and acquisitions: Their effect on the innovative performance of companies in high-tech industries. *Research policy* 35(5): 642-654
- De Man, A. P., & Duysters, G. (2005). Collaboration and innovation: a review of the effects of mergers, acquisitions and alliances on innovation. *Technovation*, 25(12), 1377-1387.
- Ferreira, M. P., Santos, J. C., de Almeida, M. I. R., & Reis, N. R. (2014). Mergers & acquisitions research: A bibliometric study of top strategy and international business journals, 1980–2010. *Journal of Business Research*, 67(12), 2550-2558.
- John, C. S., & Harrison, J. S. (1999). Manufacturing-based relatedness, synergy, and coordination. *Strategic Management Journal*, 20(2), 129-145.
- Lichtenthaler, U., & Ernst, H. (2012). The performance implications of dynamic capabilities: The case of product innovation. *Journal of Product Innovation Management*, 10.
- Makri, M., Hitt, M. A., & Lane, P. J. (2010). Complementary technologies, knowledge relatedness, and invention outcomes in high technology mergers and acquisitions. *Strategic Management Journal*, 31(6), 602-628.
- Miozzo, M., DiVito, L., & Desyllas, P. (2015). When do Acquirers Invest in the R&D Assets of Acquired Science-based Firms in Cross-border Acquisitions? The Role of Technology and Capabilities Similarity and Complementarity. *Long Range Planning*.
- Pervan, M., Višić, J., & Barnjak, K. (2015). The Impact of M&A on Company Performance: Evidence from Croatia. *Procedia Economics and Finance*, 23, 1451-1456.
- Raju, T. N. (2005). William Sealy Gosset and William A. Silverman: two “students” of science. *Pediatrics*, 116(3), 732-735.
- Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic management journal*, 18(7), 509-533.

Valentini, G., & Di Guardo, M. C. (2012). M&A and the profile of inventive activity. *Strategic Organization*, 10(4), 384-405.

Yu, Y., Umashankar, N., & Rao, V. R. (2015). Choosing the right target: Relative preferences for resource similarity and complementarity in acquisition choice. *Strategic Management Journal*.

Zhang, Y., Shang, L., Huang, L., Porter, A.L., Lu J., Zhu D. (2016). Similarity-based Patent Portfolio Analysis: Methodology and a Case Study for Technology Mergers and Acquisitions. *Journal of Informetrics*, under review.



PROPOSAL FOR ORGANISING A SPECIAL SESSION

Topic: Measuring Innovation in Government

Proposed Format: Roundtable Session (90 minutes)

Purpose: To discuss international experiences in measuring innovation in the public sector.

Novelty: Attempts to measure innovation capabilities have historically focused on measuring this in firms and/or economies. Less effort has been made to measure innovation in government.

Activities (agenda): Highly informed experts to discuss the experiences of South Korea, the UK, Europe, Nordic Countries, the OECD, Australia and the United Arab Emirates. The session will highlight best-practices and key learnings from international undertakings with regard to measuring innovation in the public sector.

Anticipated Participants: Academics, analysts and government officials with interest in promoting and measuring innovation-related activities in the public sector.

About the Organisers: INSEAD and the Mohammed Bin Rashid Centre for Government Innovation.

Chair of the Session: Dr. Sami Mahroum – Academic and Executive Director of the INSEAD Innovation and Policy Initiative, Abu Dhabi.

WHAT DO WE MEAN BY INNOVATION IN GOVERNMENT?

Over the past few decades, there has been growing interest in supporting innovation in the public sector as a means of increasing the efficiency and quality of government services [1, 2].

Generally speaking, innovation in government could be defined as “The creation and implementation of new processes, products, services and methods of delivery which result in significant improvements in outcome efficiency, effectiveness or quality” [3]. A ‘Process Innovation’ is an activity oriented mainly towards enhancing ‘efficiency’. A ‘Product or Service Innovation’ is an activity oriented mainly towards enhancing the ‘effectiveness’ of government. A ‘Policy Innovation’ is mainly oriented at enhancing outcomes. These categorisations are important for identifying and selecting the metrics of measurement and to make distinctions between inputs, outputs, and outcomes.

INNOVATION IN THE PUBLIC SECTOR VS. PRIVATE SECTOR

Attempts to measure innovation capabilities have historically focused on measuring this in firms and/or economies. Less effort has been made to measure innovation in government. The vast majority of existing approaches to measuring innovation in government are based on manuals and methodologies originally developed to measure innovation at the firm and macro-economy level, despite the fact that there are a number of ways in which the measurement of innovation in the business sector differs from the public sector. For example, innovation in business can be measured in terms of the commercial return on investment. Thus, the due diligence that precedes any investment decision about the success and the failure of innovation activity in the business sector has ultimately one indicator; which is profitability. However, in the public sector, due diligence can take the form of cost-benefit analyses in which social and environmental returns are the key indicators of success or failure.

Another important distinction is that innovation in the public sector is not driven by competition for market share, but by the necessity of creating public value. In fact, as one study of public sector innovation has shown, many public sectors innovations originate from frustration with the status quo or in response to a crisis, or even the mere inclination to do the right thing or to prevent a crisis [4]. In the private sector, increasing choices is valuable in its own right, but in public services if the extra choices are not wanted or needed, or give wider but ultimately poorer service, then innovation does not serve to improve [1].

Business sector decision-makers are ultimately accountable to the shareholders. Likewise, public sector innovators have public value and stakeholders' interests in mind when they embark on innovation activities. The differences in these accountabilities have implications for measurement. For example, while businesses will suspend an inefficient production process or a product with low return on investment, public sector organisations may justify the continued running of a financially inefficient programme on the basis of some non-commercial return on investment. In fact, in principle, one would expect government to invest only in areas where there is a clear- or suspected- market failure and a high level of expected positive externalities. Efficiency and effectiveness (return on investment) require very different performance measurement metrics when private versus public sector projects are under consideration.

There has been little research into the measurement of innovation specifically in government. This can be attributed to a traditional focus of interest in either surveying private firms (through CIS) or on indicators of innovation in an economy as a whole. The transferability of innovation metrics from private to public sectors is not however valid or practical due to differences in structure and attitudes towards risk-taking, profit motivation and adaptability to changing circumstances [5]. In fact, the public sector has often been viewed as a passive recipient of innovations from the private sector. Nowadays, there is a growing understanding that public sector innovation may have a considerable effect not only on the efficiency of public services, but also on the private sector's propensity to innovate [6]. Efforts to measure innovation in government have been undertaken in South Korea, the UK, Europe, Nordic Countries, the OECD, Australia and the United Arab Emirates.

Against this backdrop, the proposed roundtable session will provide a much-needed review of these major international efforts over the past decade to measure innovation in the public sector or to benchmark government innovation across countries. The review will highlight the objectives, unit of observation, scale, and methods used in these various initiatives, in addition to key lessons learnt.

REFERENCES

- [1] Hartley, J. (2005). Innovation in governance and public service: past and present. *Public Money & Management*, Vol. 25, no. 1, pp. 27-34.
- [2] Osborne, L. and Brown, S. (2013). *Handbook of innovation in public services*. Cheltenham: Edward Elgar Publishing Ltd.
- [3] Mulgan, G. and Albury, D. (2003). *Innovation in the public sector*. Strategy Unit, Cabinet Office, October 2003.
- [4] Walter, J. (2001). Understanding innovation: what inspires it? what makes it successful? *Governing Magazine*, December 2001.
- [5] The Partnership for Public Service and IDEO (2011). *Innovation in government*. New York: Partnership for Public Service.
- [6] Bloch, C. (2010). *Measuring public innovation in the Nordic Countries* [WWW]. Available from: www.mepin.eu



Measuring Innovation in Government: An International Review and Case Study of the UAE

Sami Mahroum*, **Yasser Al-Saleh**** and **Shatha Al-Hashmi*****

* INSEAD Innovation and Policy Initiative, Abu Dhabi.

** Mohammed Bin Rashid School of Government, Dubai.

*** Mohammed Bin Rashid Center for Government Innovation, Dubai.

ABSTRACT

Innovation is key to achieving multiple government functions and objectives, from ensuring the welfare and quality of life of citizens to advancing the economy as a whole. In the United Arab Emirates (UAE), innovation has become a strategic priority for the UAE Government, attracting substantial government investment in numerous public sector innovation initiatives. In order to monitor progress towards achieving its many innovation-driven aspirations, the UAE Government has pioneered the adoption of key performance benchmarks for monitoring innovation capability and performance. This paper examines UAE's experience in this regard as an emerging economy and draws comparisons with the experiences of other countries that have attempted at developing frameworks and indicators for innovation management in government.

1. INTRODUCTION

Around the world, there is a growing interest with regard to supporting innovation in the public sector as a means of increasing the efficiency and quality of government services [1-3]. Innovation in government may be broadly defined as “The creation and implementation of new processes, products, services and methods of delivery which result in significant improvements in outcome efficiency, effectiveness or quality” [4]. Setting an agenda for innovation measurement quite often serves the purpose of providing evidence for diagnostic purposes. Across all layers of government, decision-makers are constantly looking for information and data to enhance the quality of their policy decisions. Performance measurement has not only become a hallmark of modern government activity for monitoring and evaluation, as well as for achieving accountability and transparency, it is increasingly being used as an instrument of ‘soft power’. It is utilised to nudge key stakeholders or the population at large in a desired direction. Government officials may, for example, publicise rankings of economic competitiveness, education, health, or innovation to increase levels of awareness within local communities.

There has been little research into the measurement of innovation specifically in government [5]. This can be attributed to a traditional focus of interest in either surveying private firms – through CIS (Community Innovation Surveys) – or on developing composite indices of innovation in an economy as a whole. The most prominent examples of the latter are the Innovation Scoreboard of the EU [6]; ‘Science, Technology, and Industry Outlook’ from the OECD [7] and the ‘Nordic Innovation Monitor’ [8]; as well as indices developed by

UNCTAD [9] and the World Bank [10]. The European Commission explains that “By aggregating a number of different variables, composite indicators are able to summarise the big picture in relation to a complex issue with many dimensions” [11]. These composite indices do not, however, come without limitations; they may send misleading policy messages if they are poorly constructed or misinterpreted. They may invite simplistic policy conclusions and the selection of sub-indicators and their weighting could be the subject of political dispute [12].

Transferability of innovation metrics from the private to public sectors is not, however, valid or practical. This is due to differences in structure and attitudes towards risk-taking, profit motivation and adaptability to changing circumstances [13]. In fact, the public sector has often been viewed as a passive recipient of innovation from the private sector. There is, however, nowadays a growing understanding that public sector innovation may have a considerable effect on not only the efficiency of public services, but also on private sector propensity to innovate [14, 15]. Efforts to measure innovation in government have been undertaken in South Korea, the UK, EU, Nordic Countries, OECD, Australia and the United Arab Emirates (UAE).

The paper was developed using desk research and interviews. The paper begins by benchmarking international efforts to design and develop indicators, metrics and data to create an information infrastructure for measuring innovation in government. There is a scarcity of published data on measuring innovation in government. A research on scholar.google.com for the words “measuring innovation in government” and “innovation measurement in government” on July 9th, 2016 has yielded zero results. This is because historically the focus of scholarly work has been primarily on measuring innovation capabilities in firms and/or economies – as opposed to governments. However, in recent years there have been several instructive experiments across this world in measuring innovation in government, such as in South Korea, the UK, Australia, the Nordic Countries, EU, the OECD and most recently in the United Arab Emirates. We, thus, believe the insights generated by our research should be of tremendous interest to academics, analysts and policy officials in developing and developed countries alike.

2. INTERNATIONAL EXPERIENCE OF MEASURING INNOVATION IN GOVERNMENT

This section provides a review of major international efforts – undertaken over the past decade – to measure innovation in the public sector or to benchmark government innovation across countries. The review will highlight the objectives and methods used in these various initiatives, in addition to any key lessons learnt. Surveying the literature and the Internet, we were able to identify seven substantial efforts to measure innovation in government that have been undertaken by government or on behalf of government. These are single country exercises (Australia, Denmark, South Korea, and the UK) and multilateral exercises (EU, Nordic and the OECD). Below, we provide a brief description of each of these exercises and then we summarise shared and varied elements among these exercises.

2.1. single-country initiatives

2.1.1. Australian Public Sector Innovation Indicators (APSII)

The APSII project was established with the aim of equipping public sector organisations with data to obtain a better understanding of their innovation performance and capability. The project has been undertaken across three phases, namely:

- Development of a conceptual framework for measuring public sector innovation, completed in 2011.
- Design of a pilot ‘Australian Public Sector Innovation Survey’. Although the project was championed by the Australian Government, a technical group was formed to develop the survey comprising experts from academia and the Australian Bureau of Statistics. A pilot survey was carried out in 2012 with surveys sent to 473 individuals in 83 agencies. Of these, 367 responses were received from 61 government agencies.
- Development of Australian Public Sector Innovation Indicators.

2.1.2. South Korea

The GII was a survey initiative by the South Korean Government in which the participation of all 498 public agencies was mandatory. The aim of the survey was to gauge the level of innovation in Korean government agencies, to identify areas of weakness and develop action plans to enhance innovation capacity. The overall results of the index further served as a reference for the development of a national innovation strategy. Data was gathered through a web-based diagnostic and reporting system that worked out a weighted average of many sub-indices. The GII was organised around two main pillars, Innovation Activation (including adoption and implementation) and Diagnostics (including readiness, alignment, and internalisation)

2.1.3. The UK

Measurement of public sector innovation in the UK has been spearheaded by a quasi-non-governmental organisation – the National Endowment for Science, Technology & the Arts (NESTA)¹. The work itself was commissioned out to multiple parties, including academics. A key objective of the project was exploring various conceptual frameworks for measuring innovation in the public sector. The results of the study included development of a framework with five elements: *inputs* to innovation; innovation *processes* within the organisation; *outputs* of the innovation process; general *outcomes* of innovation; and external factors or *framework conditions* that affect innovation in public sector organisations [16]. A total of 175 public entities were interviewed by phone on a voluntary basis. The interviews included questions on capabilities such as percentage of university degree holders, innovation strategy, procurement, sources of innovation, any barriers and performance makers such as research, consultancy, design and training.

2.1.4. Denmark

The ‘Danish Innovation Barometer’ (2014) is an on-going undertaking in Denmark, the results of which are not yet available. It is based on a two-tiered survey, namely (i) ‘Highest organisational level’ – ministry or regional government level; and (ii) ‘Workplace’ units such as schools and hospitals. A small set of questions on innovation strategy were asked at the highest organisational level and a full innovation survey was distributed to the ‘workplace’

¹ Now, registered as a charity re-named ‘Nesta’.

units. A particular focus was placed on providing examples of the latest innovation in order to understand how organisations innovate. Another novel element of this survey is that is conducted at the lowest level of individual government entities, down to as few as three employees.

2.2. Multi-Country Initiatives

2.2.1. EU Innobarometer 2010 Survey (European Commission)

The 9th ‘Innobarometer on Innovation in Public Administration’ was conducted in October 2010 in the 27 Member States of the EU and Norway and Switzerland using a questionnaire modelled around the Eurostat Community Innovation Survey, which measures innovation in businesses. The primary focus of the survey was on output activity, such as the number and percentage of new services or processes introduced and on barriers to innovation.

2.2.2. Nordic Countries

The MEPIN ‘Measuring Public Innovation in Nordic countries’ project was carried out from 2008 to 2010 with the aim of developing both guidelines for data collection and a questionnaire for collecting internationally comparable data. A pilot questionnaire was developed and tested and a total of 2,012 public organisations were surveyed in the five Nordic countries. A summarised comparison of the Nordic countries with the Australian, Korean and the UK experiences is provided in Table 1.

Table 1: Summary of Key International Experiences in Measuring Innovation in Government

	Inputs	Outputs	Outcomes	Capabilities	Performance
Australia	✓ - Expenditures - HR - Technological infrastructure	✓ Tangible and intangible outputs (such as IPs and trademarks)	✓ - Social & environmental impacts - Employee satisfaction	✓ - Internal Capabilities	✓ Quality, efficiency and productivity
Korea	✓ - Innovation leadership - Vision & strategies	✗	✗	✓ - Innovation readiness - Personnel capabilities	✗
Nordic Countries	✓ - Resources (time/money) - Training	✓ - Spin offs - Licences & patents	✓ - Intermediate effects - Long-term impacts	✓ Cooperation & innovative procurement practices	✓ Driving forces & barriers of innovations
UK	✓ - Incentives - Training - ICT infrastructure	✓ Increase in outputs & pilots	✓ Impacts on the organisational outcomes	✓ Organisational enablers of innovation	✓ Improvements in efficiency & organisational performance

2.2.3. NESTI Taskforce on the Measurement of Public Sector Innovation (OECD)

An OECD working party of ‘National Experts on Science and Technology Indicators’ (NESTI) launched a task force in 2009, with the aim of developing statistical guidelines for collecting internationally-comparable data on innovation in the public sector. The task force, led by Denmark and the UK, started by producing a scoping paper that pointed out critical measurements issues and also proposed options on what guidance it should contain and a focus for debate. The paper was discussed in subsequent meetings and feedback obtained from OECD member countries.

3. MEASURING INNOVATION IN THE UAE GOVERNMENT

Interest in innovation metrics in the UAE began with the Dubai Quality Award, initiated by the Department of Economic Development in 1994. Based on the ‘Excellence Model’ created by the European Foundation for Quality Management, the aim of this trophy is to recognise and award innovative organisations from the private sector. In 1997, His Highness Sheikh Mohammed Bin Rashid Al Maktoum, UAE Vice President, Prime Minister and Ruler of Dubai made the visionary decision to establish the ‘Dubai Government Excellence Programme’ targeting public sector entities in Dubai. Recognised as the first integral programme for government excellence in the world, it has been the driving force behind the empowerment of the local public sector to provide distinctive services. Innovation was initially set as one of the categories, accounting for around 5% of the total excellence score, which in turn allocates 50% to ‘enablers’ and 50% to ‘results’. Every three years, each government entity in Dubai is asked to submit a file application, which is then closely scrutinised by a team of independent international experts using the ‘RADAR logic’ – a dynamic assessment tool that provides a structured approach to questioning the performance of an organisation. Adopting this methodology has reduced the potential for over-reporting, which is a problem experienced in some of the aforementioned international experiences. Moreover, the formal assessment is often supplemented by sudden visits by a ‘mystery shopper’ who rates the performance of government departments and their directors [17].

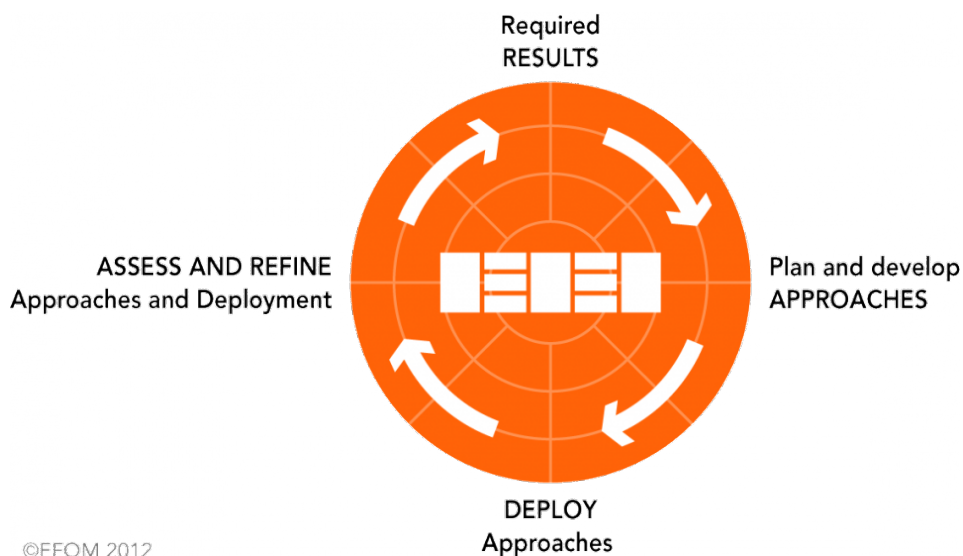


Figure 3: RADAR Logic [18]

In 2000, a new category was introduced, rewarding ‘innovating employees’ on an annual basis. An important development was the launch of the ‘IBDAA – Mohammed Bin Rashid Al Maktoum Initiative for Government Innovation’ – in 2012. The initiative set four award categories, namely; ‘Innovative Idea’, ‘Innovating Employee’, ‘Innovating Leader’ and ‘Government Entity Fostering Innovation’. It also hosts the following activities:

- Dubai Forum for Government Best Practices, an annual high-profile event held since 2008.
- Training courses for government staff, equipping them with the knowledge, methodologies and tools available for assessing quality and excellence in government. In 2013, an

‘Excellence Experts Diploma’ was launched in collaboration with the American University of Sharjah.

- One-day seminars and workshops, raising awareness of the principles and best practice in government excellence.
- Online Forums for Dubai Government Innovators – previous winners of the ‘innovating employees’ award are invited to participate in knowledge-sharing and brainstorming sessions
- An exhibition to showcase the accomplishments of winners of all award categories.
- Development of a conceptual framework for measuring excellence and innovation in the public sector – an ongoing activity undertaken by the Executive Council of Dubai.

Federal Government interest in innovation metrics began with the establishment of the Prime Minister’s Office (PMO) in 2006. The ascension to this role of Sheikh Mohammed Bin Rashid meant that the newly established PMO borrowed many of the existing practices of the Dubai Government. These included various initiatives used to track and monitor government performance along a long list of indicators, many of which are relevant to innovation. In this regard, it is not unlike the Australian case, where the Federal Government has benefitted from measurement experience in the State of Victoria. The UAE Federal Government was inspired and benefits from existing practice at a local level.

In 2009, the UAE Federal Government put in place the Sheikh Khalifa Government Excellence Programme. This constitutes a roadmap for governments seeking to reach new heights beyond excellence, to achieve performance leadership and transform into a leading, innovative and smart government. The main features of this programme are as follows:

- Focus on innovation, which is made an official category.
- Observance of variations in the work nature of government entities. In order to ensure a fair assessment, some criteria have been identified to particular entities – based on their work nature and functions.
- Focus on delivering smart services, in line with the ‘Smart Government Strategy’ and the ‘Emirates Government Service Excellence Programme’. The aim of the latter is to develop skills among customer service staff to deliver outstanding government services.
- A streamlined process for application and assessment. Applications must be submitted online and then electronically distributed to specialists who will make a field visit to the participants to compare adopted working systems with global best practices. They will also make sure that the results are linked to working systems and identify the presence of new and pioneering practices.
- Ease and clarity of criteria, with each criterion divided into two main sections: capabilities and results associated with those capabilities.

It was in 2015 that the UAE Government elevated the topic of innovation to its highest level by designating that year the ‘Year of Innovation’. The country conducted various innovation-related initiatives, and witnessed the launch of the ‘Fourth Generation of the Government Excellence System’ – a world first results-oriented system designed to develop government performance. Building on the experience of over twenty years in performance measurement, the system now recognises ‘innovation’ as a key category carrying a total weight of twenty per cent.

‘Shaping future’ (5%) focuses on developing organisational capabilities with regard to forecasting and strategic planning for the future, whereas ‘innovation management’ (15%) focuses on efforts undertaken to create an organisational culture and work environment

conducive to innovation and creativity. The underlying metrics vary according to the work nature of the entity in question, but include a survey to measure innovation readiness in the organisation, the percentage of staff trained in innovation, percentage of innovations put forward by staff, number of research papers by staff, percentage of innovations patented, number of workshops focused on innovation, percentage of budget allocated to innovation, number of new services/products, number and size of innovation projects completed in collaboration with partners.

Reflecting on the UAE experience, two main challenges have been identified; ambiguity of the concept of innovation on the part of staff and a lack of supporting data. The first challenge has been addressed through conducting numerous information sessions and awareness-building events. In terms of the second, the availability of measurement data is expected to improve following the issuance of a ‘Open Data Law’ in 2015. Furthermore, it is interesting to note that the use of performance indicators in the UAE is aimed primarily at recognition and reward and less on penalising and resource re-distribution. Occasionally, however, Sheikh Mohammed Bin Rashid adopts a ‘naming and shaming’ approach to nudge government staff to excel. For example, during the awards ceremony for the 19th edition of the Dubai Government Excellence Programme in April 2016, Sheikh Mohammed read out the names of the three worst-performing entities. He said “some officials are slacking this year and have not achieved positive results. They stopped working hard, believing that they have reached the top, which is a misconception”. Another important reason for assessing performance is monitoring and giving internal feedback. To that end, it is interesting to note that data collected in government-wide performance system called ‘Aada’ gives different levels of access to data, depending on the seniority of the user. Managers are able to benchmark the performance of their entities vis-a-vis the average and median performance of other units in government but without access to unit-level data, which is only accessible to ministers. Indicators are used to provide regular feedback and in the long-term this can have an impact on decisions relating to the leadership and staffing of low-performing agencies.

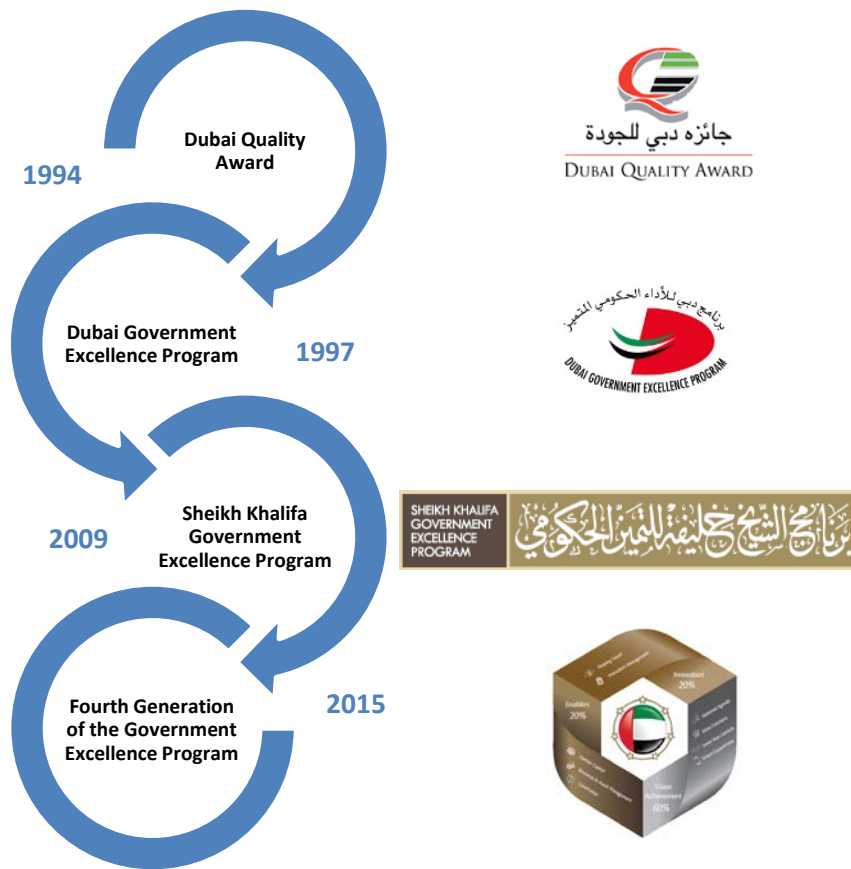


Figure 4: The Evolution of Measuring Government Innovation in the UAE

4. THE UAE VIS-A-VIS INTERNATIONAL EXPERIENCES

To conclude this paper, it is useful to reflect on how the UAE’s experience compares with other international endeavours. Table 2 lists the topics covered by key international surveys and contrasts them with those covered by the UAE Government. It can be clearly seen that the UAE has the most comprehensive scope with regard to innovation-related measurement topics. For instance, the UAE is the only country thus far to have incorporated detailed measures for governance aspects, innovation leadership, satisfaction/happiness levels of users and partners and smart service applications as part of its government excellence benchmarking programme.

Table 2: Benchmark Topics Addressed by MEPIN, EU, NESTA and the UAE

Questions/topics	MEPIN	EU Innobarometer	NESTA Public Sector Innovation Index	UAE
Types of innovations	Product (good and service), process, organisational and communication	Service, process, organisational and communication	Service, process, organisational and communication	Product, process, individual, organisational and communication
Identifying innovation	Novelty, who developed the innovation	Novelty, who developed the innovation, % services that new or significantly improved	Novelty, % service dev. That involves users, % service dev. that generated new ideas used elsewhere	Novelty, who developed the innovation, implementation and the impact
Innovation activities	Innovation expenditures (actual amounts or intervals)	Not included	Intangibles expenditures (software, training, design, consultancy, research)	Expenditures, forecasting, various innovation management activities
Procurement	External innovation purchases, innovative procurement, use in procurement activities	Procurement practices, cost vs. innovation, procurement results	Purchase of software, research, consultancy	Purchase of software, research, consultancy
Human resources	Innovation personnel	% with university degree, % in innovation project teams	% with university degree	Innovative personnel and leaders
Objectives and effects	Objectives	Effect (positive and negative), expected impact in future (2 years)	Effects	Objectives, enablers, impacts
Innovation strategy and organisation	Strategy, management and staff	Management and staff	Strategy, management and staff	Strategy, management and staff
Framework conditions	Barriers	Barriers	Framework conditions (Positive and negatives)	Enablers, framework conditions
Additional topics	Main sectors, public funding, number employees	Number employees, main sector, geographic area served	Self-evaluation against peers for performance measures	Governance, collaboration, smart applications, satisfaction levels

Whilst most of the innovation measurement exercises have been experimental in nature, it is worth noting that the UAE, South Korea and Australia have adopted their benchmarking results for use (see Table 3). Table 4 shows that the UAE and Korean Governments share further similarities with regard to the purpose/usage of their innovation benchmarks. Both of these countries use the results of their innovation measurements for the purposes of monitoring and evaluation, control and steering, agenda-setting, awareness-building and mobilisation of concern. This can be interpreted as an indication of the UAE Government's ambitious spirit to develop a vigorous science and technology policy agenda such as the one adopted by South Korea since the 1970s. Nevertheless, a distinctive feature of the UAE approach is the emphasis on rewarding top performers. Various competition-based prizes have recently been created, around the world, for the purpose of awareness-raising through sending signals about the value of certain types of behaviour or activity [19]. Holding such

performance-based competitions in the UAE's public sector has proven to be an effective motivating mechanism for civil servants to excel with the hope of securing high-profile recognitions. These awareness building efforts have also been used as a nudging lever that mobilises resources to serve particular policy goals such as innovation, smart services and quality in the public sector.

Table 3: Measurement Exercises of Government Innovations by Type

	Australia	EU	Nordic	OECD	UK	South Korea	UAE
Experimental		X	X	X	X		
Executable	X					X	X

Table 4: Usage in Countries where Measurement Results are acted upon by Government

	Australia	South Korea	UAE
Decision-making			
Monitoring and Evaluation	X	X	X
Accountability and Transparency			
Control and Steering		X	X
Awareness-building	X	X	X
Agenda-setting	X	X	X
Mobilisation of Concern	X	X	X

REFERENCES

- [1] Arundel, A., Casali, L. and Hollanders, H. (2015). *How European public sector agencies innovate: the use of bottom-up, policy-dependent and knowledge-scanning innovation methods*. *Research Policy*, Vol. 44, pp. 1271-1282.
- [2] Hartley, J. (2005). *Innovation in governance and public service: past and present*. *Public Money & Management*, Vol. 25, no. 1, pp. 27-34.
- [3] Osborne, L. and Brown, S. (2013). *Handbook of innovation in public services*. Cheltenham: Edward Elgar Publishing Ltd.
- [4] Mulgan, G. and Albury, D. (2003). *Innovation in the public sector*. Strategy Unit, Cabinet Office, October 2003.

- [5] Torugsa, N. and Arundel, A. (2015). *The nature and incidence of workgroup innovation in the Australian public sector: evidence from the Australian 2011 State of the Service Survey*. Australian Journal of Public Administration, Vol. 75, pp. 202-221.
- [6] PRO INNO Europe (2009). *European innovation scoreboard (EIS) 2009*. Brussels: European Union.
- [7] OECD (2010). *OECD Science, Technology and Industry Outlook 2010*. Paris: OECD.
- [8] Norden (2009). *Nordic Innovation Monitor*. Copenhagen: Nordic Council of Ministers.
- [9] UNCTAD (2005). *World investment report 2005*. New York: United Nations.
- [10] World Bank (2010). *World Development Indicators 2010*. Washington DC: The World Bank.
- [11] European Commission (2003). *Third European report on science and technology indicators*. Brussels: European Commission.
- [12] Nardo, M. and Saisana, M. (2008). *OECD/JRC Handbook on constructing composite indicators: putting theory into practice*. Paris: OECD.
- [13] The Partnership for Public Service and IDEO (2011). *Innovation in government*. New York: Partnership for Public Service.
- [14] Bloch, C. (2010). *Measuring public innovation in the Nordic Countries* [WWW]. Available from: www.mepin.eu
- [15] Australian Government (2011). *Measuring innovation in the public sector: a literature review*. The Australian Public Sector Innovation Indicators Project, April 2011.
- [16] Bloch, C., Jorgensen, L., Norn, M. and Vad, T. (2009). *Public sector innovation index – a diagnostic tool for measuring innovative performance and capability in public sector organisations*. London: NESTA.
- [17] Hvidt, M. (2009). *The Dubai model: An outline of key development-process elements in Dubai*. International Journal of Middle East Studies, Vol. 41, no. 3, pp. 397-418
- [18] European Foundation for Quality Management (2012). *Radar logic* [WWW]. Available from: <http://www.efqm.org/efqm-model/radar-logic>
- [19] Kouzmin, A., Loffler, E., Klages, H. and Korac-Kakabadse, N. (1999). *Benchmarking and performance measurement in public sectors: Towards learning for agency effectiveness*. The International Journal of Public Sector Management, Vol. 12, no. 2, pp. 121-144.



Baseline of indicators for R&D and Innovation in ICT: a tool for decision-making, design and monitoring of public policies¹

Henry Mora Holguín*, Diana Lucio-Arias**, Sandra Zárate***, Nayibe Castro**** and Clara Pardo*****

[*hmora@ocyt.org.co](mailto:hmora@ocyt.org.co)

Colombian Observatory of Science and Technology, Carrera 15 # 37-59, Bogotá, 111311 (Colombia)

** dlucioarias@gmail.com; szarate@ocyt.org.co; ncastro@ocyt.org.co; cpardo@ocyt.org.co

Colombian Observatory of Science and Technology, Carrera 15 # 37-59, Bogotá, 111311 (Colombia)

INTRODUCTION

Development and implementation of sophisticated strategies to improve competitiveness of sectors relies on precise monitoring of the sectors dynamics and particularly, evolution of scientific and technological development and innovation (STI) generating capacities. In a knowledge based economy, non-technological innovation plays an important due to the importance of information and knowledge management for individuals and organizations (OECD, 2011).

According to the World Economic Forum, the role of ICT in stimulating economic growth and creating new employment opportunities for highly qualified personal has never received as much attention as today and as a result it has become a common concern for researchers. ICT's positive impacts in the efficiency of firms has been widely acknowledged and allows businessmen to optimizer their firms production and mobilize resources to other more productive investments. ICTs are also regarded as an innovation source that can accelerate growth, favor technology adoption and adaptation, and promote technological change due to their effect in reducing transaction costs and minimizing the importance of geographical distance in innovation processes.

As a result of the importance of ICTs and of monitoring STI capabilities, it is necessary to have updated and relevant statistical information that facilitates the design and monitoring of public policies for the sector. In Colombia, lack of information resulted in the initiative to create a baseline of indicators to provide information on the STI activities. The set of proposed indicators should result beneficial to the academic sector, the government, the industry and society in general. We will make a brief discussion of the importance of the baseline and the methodology underlying its design and construction.

THE NEED FOR A BASELINE OF INDICATORS

Interest of MinTIC to have detailed information on the state and dynamics of STI activities motivated a to observe and monitor sectorial performance based on three activities: In first place, the formulation of research, technological development and innovation agenda for ICT requires orientations on the prospective development of the sector; in the second place, an opportune selection and construction of a set of indicators aligned with the requirement of a

¹ This work was supported by the Colombian Ministry of Information and Communication Technologies (MinTIC) and the Administrative Department of Science, Technology, and Innovation (Colciencias).

reliable picture of the sector and in the third place, definition of an action plan to close gaps in prioritized issues of the country.

From this perspective, the baseline is useful in at least two ways: first, because it is an appropriate tool to diagnose STI activities in the sector by generating information that serves policy formulation, and second, because it allows monitoring which allows to monitor the efficiency of the designed policies and their alignment to the targeted goals.

Aligned with this, the baseline provides a representation of the dynamics of ICT industries but also of other productive sectors where the ICTs have become a fundamental strategy to introduce new innovations, for example, optimize processes, among others. Time series that make part of the baseline should be periodically updated assuring updated and relevant information.

From the normative perspective, the baseline is supported in different policy documents such as the National Development plans (DNP, 2010), policies (Colciencias, 2008), economic and social policy documents (Conpes, DNP, 2009), prospective documents (DNP, 2005), etc. All of these normative documents agree on the importance of the importance of information on STI dynamics, on the requirement to consolidate information systems and statistics on STI and policy evaluation among others.

INDICATORS BASELINE ON R&D+I IN ICT

The baseline was built under the operationalization of ICT as a sector and as a transversal strategy to other industries and sectors.

Design and construction strategy

Activities behind the selection and construction of the set of indicators for the baseline:

1. Identification of STI indicators in ICT. The purpose of this activity was to establish and comprehend the way in which measurement processes have been emphasized internationally and nationally. The following activities were developed: identification and selection of relevant references and classification of indicators according to their role in the input-process-output framework and results analysis. From the activity, 1,279 international indicators were revised, and of these, 284 agreed with the purpose of the baseline and could be adopted for Colombia. In the case of national indicators, 720 agreed with the purpose of the project and could be integrated to the baseline 59 indicators.
2. Literature review to identify information requirements. Normative documents were revised to identify specific indicators for monitoring and evaluation purposes. Laws, development plans and sectorial plans were considered in this review. As a result of this activity, dispersion in the documents proposals but some convergence on the type of information required particularly in terms of infrastructure, human resources, intellectual property, investment and STI capacities.

3. Stakeholders consults, actors and stakeholders in the ICT ecosystem were consulted in order to consider their information requirements.
4. Preliminary design of the baseline's structure and content, following the three previous activities, a consolidation and depuration process was implemented. As a result, a preliminary baseline consisting of 90 indicators aggregated around three dimensions: inputs, processes and results was established. In the next section, the final structure that make up the baseline are detailed.
5. Identification and characterization of information sources; possible sources to build the indicators in the baseline were identified and an inventory of relevant sources was built. Seven national sources were identified as relevant for the proposed baseline.
6. Prioritization of the indicators; over the extended set of indicators, experts were consulted on their importance and relevance as well as the possibilities to have them measured and updated. Three criteria were selected in order to prioritize the indicators: relevance, information availability and costs associated to the construction of the indicators. In total, 53 of the 90 indicators were prioritized to be constructed.
7. Preparation of the technical specifications record of each indicator. For each of the prioritized indicators a record, with the description of the indicator, its importance, source, information processing requirements and calculation methodology was developed. This allows traceability in the calculation of the indicator and ensures replicability in the future.
8. Measurement of the baseline for 2008-2013 in order to have a first picture of STI capacities in the sector. In this stage, considering the periodicity with which the underlying information was being collected allowed to suggest the time frequency for updating the baseline. It was also in this stage when strategies for the sectorization and disaggregation of information were designed. On this stage, some compound indicators were also designed and measured as a synthesis strategy. In total 7 synthetic indicators were proposed, one for each category.
9. Socialization of results. During the process of construction, the baseline, socialization activities with key actors of the ecosystem were carried out. The purpose of this step was to validate advances, highlight the potential uses of the baseline and to promote its appropriation. Finally, the baseline was diffused through an infographic bulletin as well as a report.

Structure of the baseline

The aggregation of the 53 indicators that composed the baseline obeyed the following:

- Dimensions: general aspects that should be considered in STI measurement in ICT sector:
 - Inputs: Resources and efforts, financial and human, required for STI activities in the ICT sector.

- Process: Intermediate activities required for the efficient transformation of inputs into outputs and results.
 - Results: Measures tangible, verifiable results which are obtained from STI activities in the ICT sector.
- Categories: These are aligned with the international standards recommended for these types of activities, e.g. Frascati Manual (OCDE, 2002), Canberra (OCDE, 1995), Oslo (OCDE, EUROSTAT, 2005) and Bogotá (RICyT, OEA, COLCIENCIAS, OCYT , 2001). These are:
- Training: Supply and demand of training programs related to the ICT sector in the technical and technological levels, professional and tertiary (masters, phd, and postdoc), as well as available supports for training and personal in training process.
 - Infrastructure: Considers aspects related with the required technology for research and innovation as well as the institutional and normative architecture required.
 - Investment: Financial resources mobilized in the execution of STI activities in the ICT sector.
 - Human resources: researchers and research groups of professionals that develop STU activities in the ICT sector.
 - Management of R&D+i: considers support activities to the execution of STI in ICT sector: indicators in this category are related to relations between agents, access to tax incentives, access to sources of bibliographic information and barriers and incentives to innovation.
 - Scientific and technological production: scientific documents and publications, specialized journals and intellectual property rights in the sector.
 - Innovation results: New or significantly improved goods and services, indicators on sales from innovative products are considered in this point.

Contributions from the baseline

The LBI provides information on the country's capabilities regarding science, technology and innovation in ICT in different times. At the same time, it is a tool to assess progress over previous measurements and is also an input to propose alternative strategies to improve performance and increase the competitiveness of the ICT sector based on its scientific and technological development and innovation capabilities.

The baseline provides important information in decision-making processes and for the design of actions for government officials, academia and business, for example:

- Government:
 - Facilitates monitoring, evaluation and formulation of public policies, plan, programs and projects in the topic.
 - Coordinate agents from different spheres in the framework of their politics and programs.

- Promote ICT for development and competitiveness in the country.
- Academia:
 - Contextualize their contributions to the sectoral capacities in the country.
 - Promote research for generation of new knowledge that provides solutions from ICT.
 - Improve the educational supply of programs so that it responds to the requirements of the industry and society.
 - Get information on public support instruments for projects in these topics.
- Industry:
 - To have information on the sectorial ecosystem to consider for strategic planning exercises, possible alliances, network formation, clusters, etc.
 - Get information on public support instruments for projects in these topics.
 - Identify the supply of professionals trained on this topic.

In addition to the abovementioned contributions, the result of the actions of government, academia and private sector actors, could generate benefits to society by facilitating and promoting the efficient use of ICT, which could solve the various problems faced by cities, improve competitiveness, efficiency, sustainability and overall the levels of welfare in society.

FINAL REFLECTIONS

The baseline proposed is an innovative instrument because it aims at measuring STI capabilities in a specific sector and could be easily transferred to other productive sectors; this could contribute to the development of the country and its transition of a knowledge based society.

This first measurement exercise of STI activities in the ICT sector is, without doubt, a key tool to know where do we start from and where are we going, to take informed decisions, formulate public policy actions, plans and instruments. Its articulation into an evidenced-based policy design could generate spillovers to other sectors in the economy as a consequence of transversality of ICTs. In this sense, it would be recommended the systematic and periodic actualization of the information in the database maintaining the proposed methodologies to promote comparability and traceability.

As was previously mentioned, not all the indicators of the baseline could be built. It is worth revising the relevance of the indicators for which information was not available and to propose strategies for their construction. Finally, it is important to promote the use and appropriation of the information in the baseline by the different actors of the ecosystem (Government, academia, private sector and society in general) as a tool to contribute to the decision making processes related to the promotion of STI in the ICT sector.

REFERENCES

Departamento Administrativo de Ciencia, Tecnología e Innovación -Colciencias-. (2008). *Colombia Construye y Siembra Futuro*. Bogotá, Colombia: Colciencias.

Departamento Nacional de Planeación –DNP-. (2005). *Visión Colombia II Centenario: 2019*. Bogotá, Colombia: DNP.

Departamento Nacional de Planeación –DNP-. (2009). *Documento CONPES 3582: Política Nacional de Ciencia y Tecnología*. Bogotá, Colombia.

Departamento Nacional de Planeación –DNP-. (2010). *Plan Nacional de Desarrollo (PND) 2010-2014 Prosperidad para Todos*. Bogotá, Colombia: DNP.

OCyT. (2014a). Análisis de la literatura de indicadores de CTI en TIC nacionales e internacionales existentes. Bogotá: OCyT.

OCyT. (2014b). Identificación de Políticas de Ciencia, Tecnología e Innovación para el sector de las Tecnologías de la Información y Comunicaciones en Colombia. Bogotá: OCyT.

OCyT. (2014c). Informe “Diseño estructural para la batería de indicadores de Ciencia Tecnología e Innovación en Tecnologías de la Información y las Comunicaciones”. Bogotá: OCyT.

OCyT. (2014d). Inventario de fuentes de información identificadas. Bogotá: OCyT.

OCyT. (2014e). Informe de comparabilidad de indicadores. Bogotá: OCyT.

OCyT. (2014f). Fichas Técnicas de los Indicadores. Bogotá: OCyT.

OCyT. (2015a). Informe de cálculo de indicadores sintéticos para la LBI I+D+i de TIC. Bogotá: OCyT.

Organisation for Economic Cooperation and Development –OECD-. (2011). *Measuring Innovation: A New Perspective*. París: OECD.



Measuring originality: common patterns of invention in research and technology organizations

David Li Tang, Erica Wiseman, Tamara Keating, and Jean Archambeault

{david.tang, erica.wiseman, tamara.keating, jean.archambeault}@nrc-cnrc.gc.ca
National Research Council Canada, 1200 Montreal Road, Ottawa, Ontario, Canada

ABSTRACT

The National Research Council of Canada (NRC) co-chairs an international working group on performance benchmarking and impact assessment of Research and Technology Organizations (RTO). The Knowledge Management branch of the NRC conducted the patent analysis portion of the benchmarking study.

In this paper, we present a Weighted Originality index that can more accurately measure the spread of technological combinations in terms of hierarchical patent classifications. Using this patent indicator, we revealed a common pattern of distribution of invention originality in RTOs. Our work contributes to the methodological advancement of patent measures for the scientometric community

INTRODUCTION

The National Research Council of Canada (NRC) co-chairs an international working group on performance benchmarking and impact assessment of Research and Technology Organizations (RTOs). An RTO is a form of public research institutions which has the responsibility to build economic competitiveness and conduct research in support of government policy. The working group is currently in its third round of benchmarking with participation from six national RTOs and a seventh state RTO as an associate member. The goal of the working group is to develop a framework to assess and improve the performance of RTOs with national mandates (Leung, Bazzacco, & Jodoin, 2014).

The Knowledge Management branch of the NRC conducted the patent analysis portion of the benchmarking study. A rapid literature review¹ was performed to identify potential indicators of interest (Maddocks & Tang, 2016). Among those considered was the Originality index (Trajtenberg, Henderson, & Jaff, 1997), which is based on the premise that inventions that cite a broad technological range of prior art are more original (Squicciarini, Dernis, & Criscuolo, 2013). To address limitations that were discovered with regards to varying levels of International Patent Classification (IPC) codes², we modified the original formula. This paper reports on the use of a modified Originality index in the benchmarking exercise, explains the rationale behind the adaptation, and discusses its use in relation to patent assessment methods in the literature.

¹ Rapid review is a valid method for scanning major research streams on a focused subject (Grant & Booth, 2009).

² IPCs can be broken down to five levels including section, class, subclass, main group and subgroup.

BACKGROUND

The importance of patent citation statistics is well established in the literature. Reviews (Maddocks & Tang, 2016; Reitzig, 2004) confirm that patent citations are more frequently used than other patent indicators to measure the value of inventions and to support technology identification and forecasting (Albert, M.B., Avery, D., Narin, F., McAllister, 1991; Carpenter, Narin, & Woolf, 1981; Hall & Trajtenberg, 2004; Shane, 2001; Verspagen, 2007). Most of the time, citation indicators are simply counts and serve as a measure of volume, however, two indicators, specifically the Originality index and Generality index measure breadth and diversity (Henderson, Jaffe, & Trajtenberg, 1998; Trajtenberg et al., 1997). The Originality index is based on the assumption that inventions that rely on a large number of diverse knowledge sources (i.e., backward citations) lead to more original results, whereas the Generality index assumes that inventions that receive forward citations from a broad range of patents are more foundational to subsequent inventions. Both indicators are based on the notion that innovation is a combinatorial process (Arthur, 2007).

PURPOSE OF THE STUDY

The purpose of this study is to contribute to the advancement of patent indicators for measuring and comparing the originality of inventions in RTOs. The scan of the literature did not identify any study that used patent indicators to specifically assess the originality of RTO inventions.³ To this end, the Originality index is an appropriate indicator for understanding RTO inventions in terms of the extent to which citations to prior art reflect the originality of patented inventions. This paper does not aim to rank specific RTOs, but instead focuses on illustrating the methodological features of the proposed indicator and how it may be used.

WEIGHTED ORIGINALITY INDEX

The Originality index was proposed by Trajtenberg et al. (1997) to measure the breadth of the knowledge that an invention draws from. The underlying assumption is that synthesis of widely divergent ideas is characteristic of research output that is highly original. This backward-looking indicator seeks to reveal the technological diversity, or breadth of the existing knowledge, on which an invention builds. This technological diversity is represented by the quantity and the spread of patent classification codes that are found on the cited patents of a given invention. The more a patent cites other patents with a greater number and broader spread of classification codes, the higher the Originality score (Hall, Jaffe, & Trajtenberg, 2001).

CHALLENGE

Given a set of citations, their total number and spread of unique classification codes has a direct impact on the Originality index. In its original formula, the calculation of originality requires a decision regarding the level of the IPC at which cited patents are counted. Changing the level (e.g., from IPC class to subclass) will not only lead to different originality results for the same invention, but could also make some inventions appear more original than

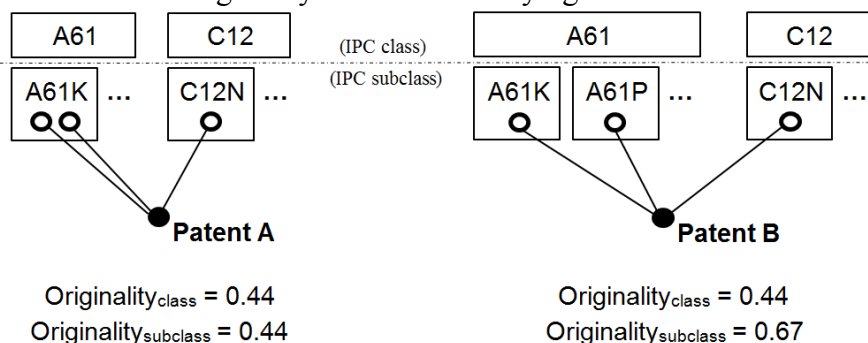
³ The literature scan also found no studies that used patent indicators to specifically assess the “novelty” of RTO inventions. In the literature, the terms “originality” and “novelty” are often used interchangeably, because the term “novelty”, as mentioned in patent assessment literature, essentially refers to the breadth and distance in the citation of technologies and this is what the Originality index is measuring.

others that were on par at a previous classification level. Tang and Wiseman (2016) quantitatively examined the impact that changing the IPC level has on the originality score and found that the difference in the total number of unique codes between any two levels leads to a significant change in originality scores. Therefore, when the originality index is used for performance benchmarking purposes, the impact due to the choice of level at which the originality score is calculated should be carefully considered in order to suit the organizational evaluation objectives.

To illustrate the impact of changing the patent classification level in the formula,

Illustration presents two inventions with the same number of cited patents which have the same number of discrete IPC *classes* but a different number of discrete IPC *subclasses*.

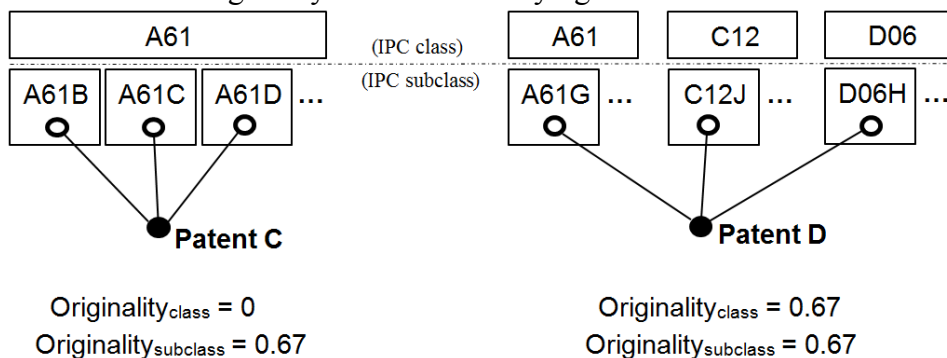
Illustration 1. Originality scores from varying IPC levels



At the *class* level, both patents have two discrete IPCs, however, if we take into account the *subclass* level, patent A has two discrete IPCs (A61K, C12N) whereas patent B has three discrete IPCs (A61K, A61P, C12N).

Similarly, Illustration 2 presents two inventions with the same number of cited patents which have a different number of discrete IPC *classes* but the same number of discrete IPC *subclasses*.

Illustration 2. Originality scores from varying IPC levels



At the *class* level, patent C has one discrete IPC, whereas patent D has three discrete IPCs, but at the *subclass* level, both patents have three discrete IPCs.

These distinctions result in the two inventions, in both scenarios, having different Originality scores depending on the level that is used in the calculation. Furthermore, it is important to

recognize that the results may differ when taking *either* the class level *or* the subclass level into account. Thus a better solution is to take *each* level into account. When each level is taken into account, it is possible to quantitatively and more comprehensively reflect the ‘distance’ between each code on a single patent. This ‘distance’ adds another dimension to the measurement of originality by reflecting the technological spread in the combination of individual codes at *all* levels. When an invention integrates technological functionalities drawn from distant domains, that invention can be considered more “novel” (Youn, Strumsky, Bettencourt, & Lobo, 2015).

SOLUTION

To account for originality for each level, Tang & Wiseman (2016) proposed a Weighted Originality index as:

$$1 - \sum_l^{m_l} W_l \sum_j^{n_l} S_{ij}^2$$

where S_{ij} denotes the percentage of citations made by invention i that belongs to technology classification j at a given level and W denotes the percentage weight assigned to each specific level l considered in the technology classification system, with all weights adding up to 1. The weighted formula essentially calculates the Originality index at three IPC levels and then sums the weighted score of each level. For example, patent D cites three patents in three *classes* and is intuitively more original than patents B and C which, although they also cite three patents in three *subclasses*, these citations come from only two *classes*. The result is a more flexible and accurate implementation of the established originality concept.

METHODS

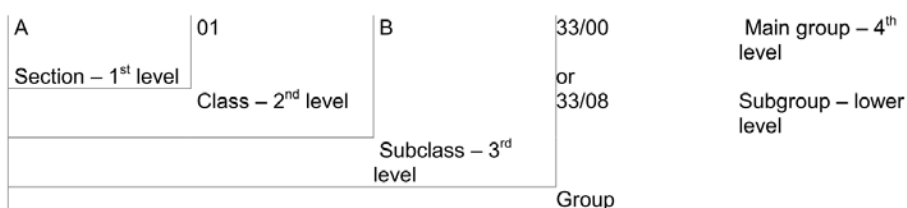
The Weighted Originality index was tested in measuring the originality of the RTO benchmarking groups’ patent portfolios, with four presented here for illustration purposes.

MEASURING THE ORIGINALITY OF RTO INVENTIONS

Patent families for each of the four RTOs were searched in a proprietary patent database within a three-year window (earliest priority date: 2011-13). A patent family is a set of patents issued in various countries to protect a single invention. Patent families, as opposed to individual patents, were chosen to ensure a consolidated list of citations (and IPCs) for a single invention to account for regional variation in disclosure obligations and examination procedures which impact the number of citations in a patent (OECD, 2009).⁴

A relational database was created to operationalize the Weighted Originality formula and to calculate scores for each invention. IPC codes, as shown below, are hierarchical and structured from general topics to more specific technologies as you expand the code.

⁴ The patent database we searched aggregates filings from 100 patent authorities at the time of this paper.



Source: IPC Guide,

WIPO, 2015

Because each level in the IPC system has a particular function, it is inappropriate to fully discount any one level in favour of another. To take each level into account, three weight schemes were tested, each emphasizing a level (class, subclass, and group) with an 85%, 10% or 5% weight. The effect of the weight parameter was analysed by calculating the arithmetic mean and median⁵ for each RTO in each of the three schemes and the results for all the RTOs were plotted on a single chart. The chart revealed two types of change in mean and median originality scores, one within each RTO and another between them across the three schemes. Finally, a weighting scheme of 10% on class, 85% on subclass and 5% on group was chosen because it is at the subclass level (four-digit IPC code) that the most meaningful differences in terms of identifying or comparing different domains occurs. For example, there is much greater differentiation between patents in A01B (Soil working; Parts, details or accessories of agricultural machines or implements) and A01C (Planting; Sowing; Fertilising) than at the more generic A01 (Agriculture; Forestry ...). Similarly, there is often less significant differences at the group or subgroup levels, for example:

- A01B 33/02: Tilling implement with rotary tools on horizontal shaft transverse to direction of travel
- A01B 33/04: Tilling implement with rotary tools on horizontal shaft parallel to direction of travel

Finally, given the weighted originality scores, a subset of highly original patents was used to identify the research domains in which each RTO had the greatest number of original inventions. An originality score of 0.72 was used as a cut off to identify the highly original inventions, here defined as patents with more technologically varied citations (in terms of IPC codes). The cut-off point roughly corresponds to the top quartile of the range of possible Originality index values [0.025, 1). This highly original set of inventions helps to characterize the research domains in which an RTO has advanced combinatorial capabilities.

RESULTS

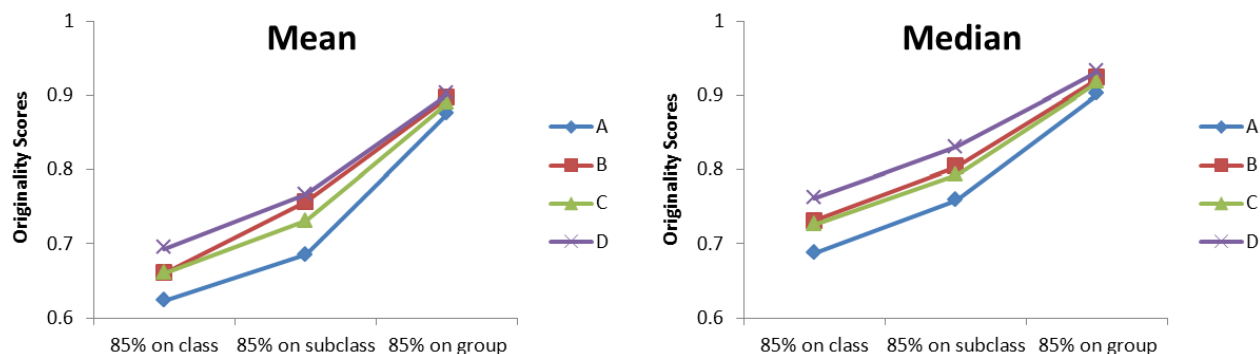
The total number of patent families of each RTO ranged from 186 to 2,300. Inventions with greater than a zero originality score were included in the analysis.

Figure 1 shows that shifting the weight from class to the more granular subclass and group levels caused an increase in overall originality scores. This result is in line with an OECD study in which the Originality index was calculated separately with four-digit and seven-digit codes (Squicciarini et al., 2013). While the general up-shift in average originality scores was expected at less aggregated levels, it is more interesting to see changes in the relative positions of RTOs. In comparing mean values, RTO B has nearly the same originality score as RTO C when calculated at the 85% weighting on class, but visibly higher mean originality

⁵ Both were calculated because inventions are not evenly distributed across the range of originality scale.

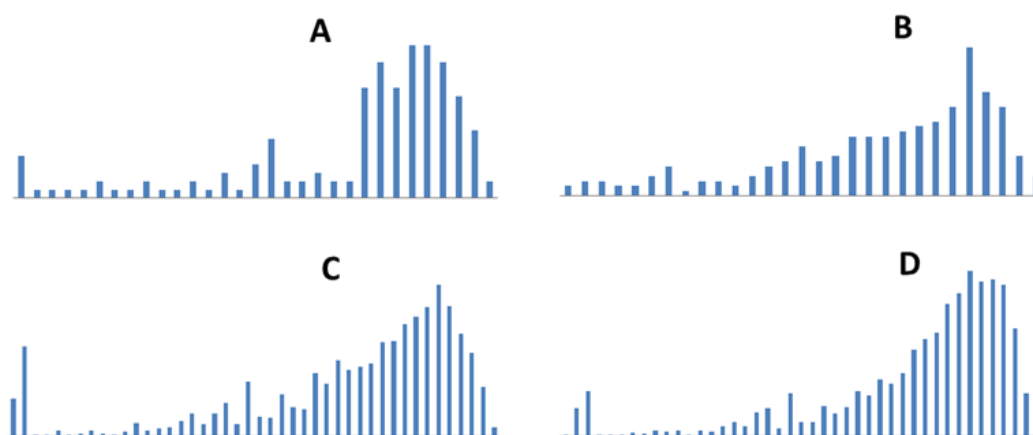
scores when calculated at the 85% weighting on subclass and is significantly closer in mean originality score to RTO D when calculated at the 85% weighting on group. These shifts, although less notable, can also be seen in the median values. Furthermore, the differences in both the mean and median originality scores for all RTOs become less distinguishable at the highest granularity level of 85% on group.

Figure 1: Shift in the average scores of four RTOs' Originality Indices.



Our major finding is concerned with commonalities in the pattern of originality in the RTOs' inventions. Figure 2 clearly shows that regardless of the volume of patent families, all RTOs demonstrate a similar distribution pattern: a small bulk of patents with a low originality score, followed by a dip (tail) and then an ever increasing number of patents with higher originality scores toward the middle (fin), with a subsequent a hump at the highest originality mark (head), ending with a decline thereafter (nose). We call it a “whale” distribution due to its shape.

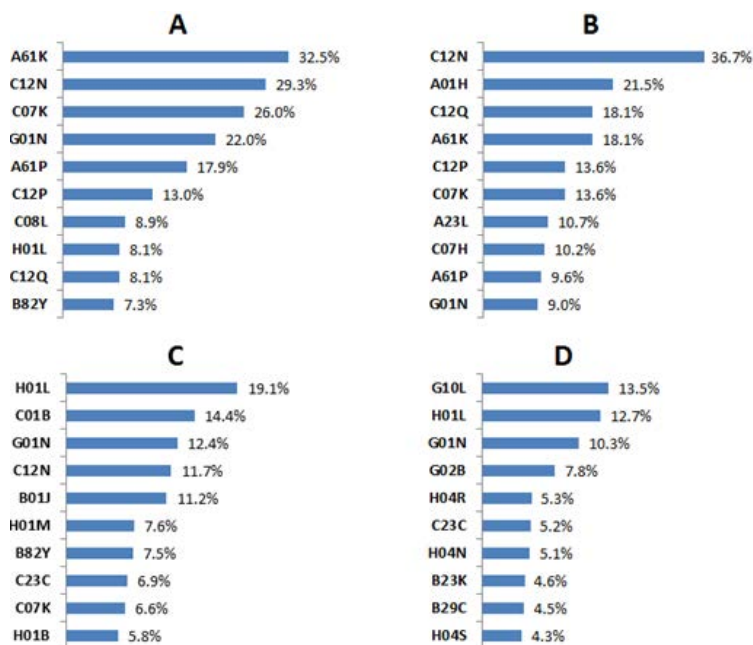
Figure 2: Distributions of RTO inventions per the Originality index.



All of the IPC classes of the patents with “high-originality” scores (defined as >0.72) were identified and the classes were then ranked by the number of inventions. Figure 3 shows that 32.5% of RTO A's high-originality patents are classified as A61K. In fact, each RTO has a unique ranking of high-originality classes, e.g., RTO A's high-originality patents appear

mainly in A61K followed by C12N, whereas RTO C's original patents appear mainly in H01L followed by C01B.

Figure 3: Concentrations of most original inventions per RTO.



As Figure 3 shows, the Originality index can be used to stratify an RTO's patent portfolio and allows for further analysis, such as identifying a signature of the IPC ranking of the high-originality inventions of each RTO.

DISCUSSION

The change in RTO's originality scores, as a result of applying the weighting scheme, provides empirical evidence to support our proposal that a weight parameter should be moderated in measuring the dispersion of patent citations. This confirms the value of the Weighted Originality index, and thus presents a methodological improvement on the original measure.

In interpreting results on RTOs' originality concentration, we have to point out that the indicator shares the inherent limitations of patent citation measures, namely:

- The higher variety of IPC codes that are required to precisely identify certain technologies
- The lack of consideration of citations to scientific literature
- The different national citation practices (e.g., U.S.A. vs Europe)
- The differences between patent offices in their use of classification codes
- The limited patent citations by inventions that rely to a much less extent on a combinatorial process of innovation (i.e., ground-breaking inventions may not have many citations)

Despite these limitations, the enhanced formula and the method that was applied here are generalizable in a number of ways. First, our relational database was also used to apply the weighting strategy on the Generality index, where forward citations, in place of backward citations are used in constructing the indicator. The Weighted Generality index was similarly found to be more flexible than the original. Second, citation measures are key components in composite indicators of innovation activity in companies, industry, and various public research institutions (e.g., Lanjouw & Schankerman, 2004; Sapsalis, van Pottelsberghe de la Potterie, & Navon, 2006). Our new formula, although only a modest and focused improvement, can be a useful building block in advancing composite indicators of innovation.

To the best of our knowledge, this study is the first to report on an empirically identified, typical “whale” shaped originality distribution of RTO inventions. Because these RTOs are located in several countries and continents, the pattern is likely representative of originality distributions across RTOs’ patent portfolios. Moreover, the method of ranking high-originality classes demonstrates an aspect of innovation and productivity that could not be shown as effectively with volume based citation measures. A first step for future research is to test if the common pattern of originality distribution (i.e., the “whale”) is also found in other research oriented institutes (e.g., corporations, academia) and if it is applicable even to patent portfolios of entire countries. Additionally future work could explore whether the RTO’s ranking of the high-originality inventions is a reflection of its innovation strategy and objectives.

CONCLUSION

To help assess patent originality, we presented a way to more accurately measure the spread of technological combinations in terms of hierarchical patent classifications. The Weighted Originality index reveals a common pattern of distribution of invention originality in RTOs which are one of the most research-intensive and innovation-oriented organizations in the world. Our work contributes to the methodological advancement of patent measures for the scientometric community.

REFERENCES

- Albert, M.B., Avery, D., Narin, F., McAllister, P. (1991). Direct validation of citation counts as indicators of industrially important patents. *Research Policy*, 20, 251–259.
- Arthur, W. B. (2007). The structure of invention. *Research Policy*, 36(2), 274–287.
- Carpenter, M. P., Narin, F., & Woolf, P. (1981). Citation rates to technologically important patents. *World Patent Information*, 3(4), 160–163.
- Grant, M. J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, 26(2), 91–108.
- Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). The NBER patent citations data file: lessons, insights and methodological tools. NBER Working Paper #8498, (3094), 1–74.
- Hall, B. H., & Trajtenberg, M. (2004). Uncovering GPTS with patent data. NBER Working Paper Series, (November).

- Henderson, R., Jaffe, A. B., & Trajtenberg, M. (1998). Universities as a source of commercial technology: a detailed analysis of university patenting, 1965–1988. *Review of Economics and Statistics*, 80(1), 119–127.
- Lanjouw, J. O., & Schankerman, M. (2004). Patent quality and research productivity: measuring innovation with multiple indicators. *Economic Journal*, 114(495), 441–465.
- Leung, F., Bazzacco, M., & Jodoin, C. (2014). Measuring performance of research and technology organisations: drivers and challenges, today and in the future. In *R&D Management Conference* (p. 911).
- Maddocks, E., & Tang, D. L. (2016). Citation metrics for benchmarking innovation (Unpublished manuscript).
- OECD. (2009). The use and analysis of citations in patents. In *OECD Patent Statistics Manual* (p. 105).
- Reitzig, M. (2004). Improving patent valuations for management purposes: validating new indicators by analyzing application rationales. *Research Policy*, 33(6-7), 939–957.
- Sapsalis, E., van Pottelsberghe de la Potterie, B., & Navon, R. (2006). Academic versus industry patenting: an in-depth analysis of what determines patent value. *Research Policy*, 35(10), 1631–1645.
- Shane, S. (2001). Technological opportunities and new firm creation. *Management Science*, 47(2), 205–220.
- Squicciarini, M., Dernis, H., & Criscuolo, C. (2013). Measuring patent quality: indicators of technological and economic value. *OECD Science, Technology and Industry Working Papers*, (03), 70.
- Tang, D. L., & Wiseman, E. (2016). Measuring inventions for organizational evaluation: a weighted approach to the Originality index, Manuscript submitted for publication.
- Trajtenberg, M., Henderson, R., & Jaff, A. (1997). University versus corporate patents: a window on the basicness of invention. *Opa*, 5, 19–50.
- Tsertsvadze, A., Chen, Y. F., Moher, D., Sutcliffe, P., & McCarthy, N. (2015). How to conduct systematic reviews more expeditiously? *Systematic Reviews*, 4(1), 160.
- Verspagen, B. (2007). Mapping technology trajectories as patent citation networks: a study on the history of fuel cell research. *Advances*, 10(1), 93–115.
- Youn, H., Strumsky, D., Bettencourt, L. M. A., & Lobo, J. (2015). Invention as a combinatorial process: evidence from US patents. *Journal of the Royal Society, Interface / the Royal Society*, 12(106).



Linking international trademark databases to inform IP research and policy¹

Steve Petrie *

* spetrie@swin.edu.au

Centre for Transformative Innovation, Swinburne University of Technology, John St, Hawthorn, VIC, 3122 (Australia)

ABSTRACT

Researchers and policy makers are concerned with many international issues regarding trademarks, such as trademark squatting, cluttering, and dilution. Trademark application data can provide an evidence base to inform government policy regarding these issues, and can also produce quantitative insights into economic trends and brand dynamics. Currently, national trademark databases can provide insight into economic and brand dynamics at the national level, but gaining such insight at an international level is more difficult due to a lack of internationally linked trademark data. We are in the process of building a harmonised international trademark database (the “Patstat of trademarks”), in which equivalent trademarks have been identified across national offices.

We have developed a pilot database that incorporates 6.4 million U.S., 1.3 million Australian, and 0.5 million New Zealand trademark applications, spanning over 100 years. The database will be extended to incorporate trademark data from other participating intellectual property (IP) offices as they join the project. Confirmed partners include the United Kingdom, WIPO, and OHIM. We will continue to expand the scope of the project, and intend to include many more IP offices from around the world.

In addition to building the pilot database, we have developed a linking algorithm that identifies equivalent trademarks (TMs) across the three jurisdictions. The algorithm can currently be applied to all applications that contain TM text; i.e. around 96% of all applications. In its current state, the algorithm successfully identifies ~ 97% of equivalent TMs that are known to be linked a priori, as they have shared international registration number through the Madrid protocol.

When complete, the internationally linked trademark database will be a valuable resource for researchers and policy-makers in fields such as econometrics, intellectual property rights, and brand policy.

INTRODUCTION

Researchers and policy makers are concerned with many international issues regarding trademarks, such as trademark squatting, cluttering, and dilution. Trademark application data can provide an evidence base to inform government policy regarding these issues, and can also produce quantitative insights into economic trends and brand dynamics (see Schautschick & Greenhalgh, 2016 for a review). Currently, national trademark databases can provide insight into economic and brand dynamics at the national level, but gaining such insight at an international level is more difficult due to a lack of internationally linked trademark data. We are in the process of building a harmonised international trademark database (the “Patstat of

¹ This work was supported by IP Australia, a listed entity within the Australian Government’s Department of Industry, Innovation and Science

trademarks”), in which equivalent trademarks have been identified across national offices. This will provide a quantitative evidence base for economics research, brand research, and government policy.

We have developed a pilot database that incorporates 6.4 million U.S., 1.3 million Australian, and 0.5 million New Zealand trademark applications, spanning over 100 years. The database will be extended to incorporate trademark data from other participating intellectual property (IP) offices as they join the project. Confirmed partners include the United Kingdom, WIPO, and OHIM. We will continue to expand the scope of the project, and intend to include many more IP offices from around the world.

A linking algorithm for the identification of equivalent trademarks

In addition to building the pilot database, we have developed a linking algorithm that identifies equivalent trademarks (TMs) across the three jurisdictions. The algorithm can currently be applied to all applications that contain TM text; i.e. around 96% of all applications.

The algorithm uses trademark application data – such as trademark text, filing date, and Nice classification – to categorise applications into groups that are likely to include equivalent TMs. A blocking (or “binning”) procedure is used in conjunction with hash tables to efficiently group TM applications by similar TM text (e.g. see Table 1). This reduces the number of pair-wise TM-TM links that need to be processed and thereby improves the efficiency of the algorithm, while at the same time reducing the likelihood of attaining false-positive links. We briefly describe the binning procedure below.

Table 1. An example of a linked bin (for “NATURALINSTINCT”), in which 10 applications (rows) submitted to the US (USPTO), Australian (IPA), and New Zealand (IPONZ) offices have been grouped together by the linking algorithm.

Office	Filing date	TM text	Nice classes
USPTO	2011-03-09	NATURAL INSTINCT	31
IPA	2008-01-09	NI NATURAL INSTINCT	3
IPA	2010-05-11	NI NATURAL INSTINCT	3, 5
IPA	2010-05-11	NATURAL INSTINCT	3, 5
IPA	2012-01-18	NI NATURAL INSTINCT	3, 5
IPA	2012-10-18	NATURAL INSTINCT	3
IPONZ	2006-08-18	Natural Instinct	5
IPONZ	2009-07-23	Natural Instinct	31
IPONZ	2013-02-14	NATURAL INSTINCT	3
IPONZ	2015-08-24	NATURAL INSTINCT	3

If we were to group TM applications into candidate linked bins under a simple (yet stringent) requirement that they have exactly identical TM text, the algorithm would fail to identify a number of TM-TM links. This is because different offices record TM text in slightly different ways. For example, the unlauded “ä” in the TM text “HB Hofbräuhaus” might be recorded as “HB Hofbrauhaus” in Australia and “Hofbr?uhaus” in the US; i.e. the leading acronym “HB” has been removed from the US TM text and the “ä” has been replaced by a question mark (“?”). Indeed, special characters such as unlauded letters, accented letters, subscripted

numbers, and superscripted numbers are often replaced by question marks in the US data, but this does not occur in the Australian or New Zealand data.

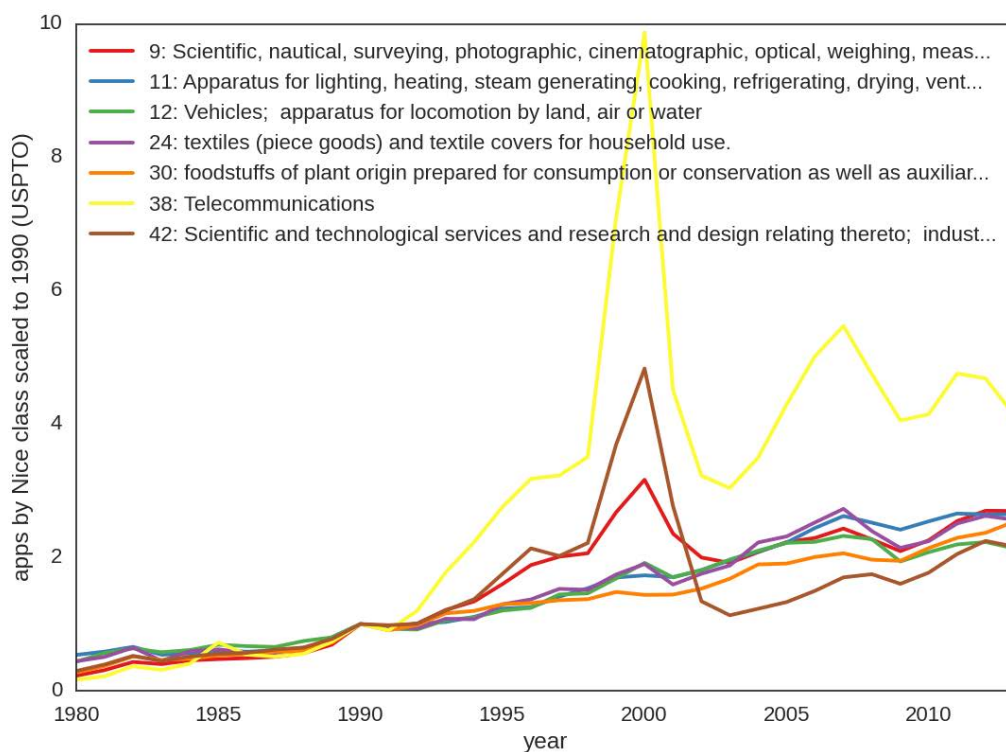
To account for potential inter-office variations in TM text, we convert all TM text to upper case and remove all whitespace and most punctuation before performing the binning procedure. Any acronyms that are either prepended or appended to the TM text are identified and removed. Question marks in the US data are handled by firstly generating bins for Australian and New Zealand data prior to US data. Then, when the US data are processed, any question marks in a given application are iteratively replaced by characters that are likely to have been substituted by the question mark, such as vowels and numbers (e.g. “HOFBRAUHAUS”, “HOFBREUHAUS”, “HOFBR1UHAUS”, “HOFBR2UHAUS”, etc...), or are simply removed (e.g. “HOFBRUHAUS”). The application is then grouped with any pre-existing bins with matching TM text. If there are no pre-existing matched bins, a new bin is created with the question mark present in the text (e.g. “HOFBR?UHAUS”).

Bins are then further divided based on filing date and Nice classification. For a given bin, if any applications are separated by a gap of > 4 years, then that grouped bin is split into separate bins along those gaps. For Nice classification, each application is simply removed from its bin if none of its associated Nice classification numbers are shared by any other application in the same bin.

The linking algorithm can potentially be adjusted to address different research questions by re-running the algorithm with different binning criteria. For instance, if a researcher is investigating how firms evolve their branding of individual products over time, then they may prefer temporally distant TM applications to remain together in the same grouped bin, rather than being separated into different bins.

In addition to investigating trends in individual firms over time, the database could also be used to investigate large-scale economic or IP trends. For example, examining the number of USPTO applications over time by Nice classification number (Figure 1) shows evidence of the dot-com bubble in the late ,90s in science- and technology-related classes (9, 38, 42), and a reduction in trademarking across all examined classes during the global financial crisis. Similar trends are evident in both IPA and IPONZ applications (data not shown).

Figure 1: Number of USPTO applications by Nice classification number, scaled to the number of applications in 1990 (which lies in a relatively stable trademarking period), for several science- and technology-related classes (9, 38, 42) as well as classes that have relatively little relation to science and technology (11, 12, 24, 30).



Current algorithm performance

The algorithm's performance can be tested using equivalent TMs that are known to be linked *a priori*, due to shared international registration number through the Madrid protocol. The algorithm successfully identifies ~ 97% of these “Known Links”. Additionally, the number of applications within the candidate positive links identified by the algorithm (1.04Mn US, 534k Australian, and 287k NZ applications) is much larger than the number of applications within the *a priori* Known Links (39k US, 65k Australian, and 35k NZ applications).

Current estimates indicate that approximately 40% of candidate positive links identified by the algorithm are false positives. However, we expect the proportion of false positives to become far smaller as we continue to improve the linking algorithm by, for example, incorporating additional application data such as firm ownership and TM images. Incorporating TM image matching into the algorithm will also make it possible to apply the linking algorithm to the remaining 4% of applications that do not have TM text data.

Using machine learning to improve the linking algorithm

A major part of improving the trademark linking algorithm will involve combining it with a separate machine learning algorithm that we have recently developed. The machine learning algorithm includes an image classification neural network (Krizhevsky, Sutskever & Hinton, 2012) that we adapted to match and disambiguate inventor names in patent records, and has

been shown to have very low false positive and false negative error rates. We briefly present the machine learning algorithm within the context of its original purpose of disambiguating inventor names.

The machine learning algorithm involves a novel matching technique whereby each pair of inventor records is compared by firstly converting the two records from raw text into an abstract visual representation, which we refer to as a “comparison image”. The neural network is able to learn important features within comparison images that indicate whether the two inventor records are likely to be a match (both inventor records refer to the same inventor) or non-match (records refer to different inventors). This is done by training the neural network on data that has been manually labelled as match/non-match. Tests on a sub-sample of the labelled data, which was withheld from the neural network during training, indicate that the algorithm can generate error rates as low as $\sim 1\%$. Given these very low error rates, we intend to further develop the machine learning algorithm and adapt it for use with trademarks.

When complete, the internationally linked trademark database will be a valuable resource for researchers and policy-makers in fields such as econometrics, intellectual property rights, and brand policy.

REFERENCES

- Krizhevsky, A., Sutskever, I. & Hinton, G. (2012), ImageNet Classification with Deep Convolutional Neural Networks, *Advances in Neural Information Processing Systems*, 25: 1097-1105
- Schautschick, P. & Greenhalgh, C. (2016), Empirical studies of trade marks – the existing economic literature, *Economics of Innovation and New Technology*, 25(4): 358-390



Detecting emerging trends and country specializations in Energy Efficiency¹.

Daniela De Filippo* and Andres Pandiella-Dominique **

[*dfilippo@bib.uc3m.es](mailto:dfilippo@bib.uc3m.es); [**apandiel@bib.uc3m.es](mailto:apandiel@bib.uc3m.es)

Laboratory for Metric Information Studies (LEMI); Research Institute for Higher Education and Science (INAECU), Carlos III University of Madrid, Calle Madrid, N. 126, Getafe-Madrid, 28038 (Spain)

INTRODUCTION

The sector of Energy Efficiency is key in the development of a country. Energy saving means cost savings and this affects GDP growth and employment. In addition, there are environmental benefits such as savings in the use of natural resources or reducing carbon emissions. If we also take into account the deduction of foreign energy dependence, it is noted that it is a sector of vital economic and strategic importance. Proof of this is that energy efficiency is a priority policy in science and technology agenda of most countries (OECD, 2011).

Although it is an area of global concern, the particularities of each context make both the issues and the areas of scientific and technological activity can be considerably different from one country to another.

To study these peculiarities, scientific publications in international databases were analyzed. These outputs are a good indicator of the intensity and specialization of scientific and technological activity of countries. Also, the characteristics of energy efficiency sector -such as the crosscutting and interdisciplinarity- make it an attractive object to be approached from a bibliometric perspective.

This methodological approach was used because bibliometrics has become one of the central models to analyze measure and evaluate different aspects of the scientific activity (Callon, Courtial, & Penan, 1995). Among its most frequent uses are, the analysis of scientific productivity (researchers, groups, institutions, disciplines or countries), and the study of its impact on the scientific community. However, it has had more innovative applications, as the detection of new research fronts, the study of collaborative networking between different institutional actors and the identification of niche of research.

Given these considerations, the aim of this work is to show the results of a study in which we analyze scientific activity in energy efficiency through bibliometric indicators. The bibliometric research, constructed upon publication data, enable us to obtain an interesting perspective of the state of scientific research in energy efficiency. This approach allow us identifying topics of interest, detecting emerging research trends and specialization of countries and determining different profile of by country. To achieve these goals, we propose the use of gross quantitative indicators, bibliometric techniques based on multivariate analysis and social network analysis.

The gross quantitative indicators allow analyzing the overall growth of publications, the scientific production by country and the thematic specialization of the area. This technique

¹This work was supported by the projects: “New horizons in research and innovation for sustainable energy and transport in the urban environment” (REF: CSO2014-51916-C2-1-R) and “Detection of new front of research and innovation in energy efficiency in Spain. Analysis of the flow of knowledge between science, industry and society” (REF: CSO2014-58889-JIN) supported by Ministry of Economy and Competitiveness.

has-been extensively validated in studies at macro and meso level as the work of Fernandez and others related to the thematic specialization of Latin American countries (Fernandez et al, 2005). Other studies using gross quantitative indicators are related to the production and specialization of African countries (Comfraria and Godinho, 2015) or focused on the analysis of publications from highly productive core countries (Bornman and Leydesdorff, 2015). Meanwhile, social network analysis has been widely used over the years for visualizing areas of knowledge and detection of emerging research areas. Borner, Chen, and Boyack (2003) reviewed techniques to map knowledge domain structures of scientific disciplines. Calero-Medina and Noyons (2008) studied an approach to reveal the structure of “Absostive capacity” field by combining mapping and network analysis techniques. Chen (2006) showed a new approach to detecting and visualizing emerging trends and transient patterns in scientific literature. Vargas-Quesada and de Moya-Anegon (2007) proposed a methodology for visualizing large scientific domains. They used approaches such as domain analysis, social networks, cluster analysis and pathfinder networks to create scientograms. In this work the social network analysis is used to determine the structure of the field. Finally, correspondence analysis (CA) is used to identify emerging areas of research and thematic country profiles. This technique has already been applied previously in bibliometric analysis (Dore & Ojasoo, 2001; Dore et al., 1996). Bordons, Bravo, and Barrigon (2004) applied CA to understand the dynamics of discipline and to identify emerging topics of a therapeutic drug (Aspirin). Lascurain, Madera-Jaramillo, Ortoll, and Sanz Casado (2010) tracked agents innovation profiles and patent subjects over time. Mauleon and Bordons (2014) identified the relationship among the sex of the inventor, the institutional sector of the assignee and the technological field of the patent is explored through correspondence analysis.

SOURCE AND METHODOLOGY

As a source of information, publications indexed in the Web of Science during the period 2000-2014 was used. The search strategy has considered the terms "Energy efficiency", "Energy saving" and "Energy storage"(and its derivatives phrases and truncation) in the field TOPICS (abstract, title and keyword). We have considered these terms after reviewing keywords used in this subject in European projects and after reviewing of broader strategies. After evaluating the adequacy of the results, 8 thematic categories (approximately 1200 documents) with results not related to energy efficiency were excluded. Once selected relevant publications, the study was conducted in four phases.

1- Getting bibliometric indicators:

After collecting publications, a relational database has been built. The information obtained has been refined and standardized and subsequently activity indicators were obtained: Annual evolution of the number of documents, publications by country, thematic production (categories WoS) and keywords.

2- Revealing the thematic structure of knowledge domain

In order to have an overview of the structure of energy efficiency field, we mapped the keywords (author keywords and keyword-plus) of the publications. We started by working with 39948 keywords. After normalization we selected 149 most frequent keywords (with 70 or more occurrences). We used a bibliometric mapping method based on keyword co-occurrences (co-word analysis). VOSviewer software were used to obtain the main clusters and to visualize the co-word network.

3- Identification emerging areas of research:

To detect new areas of research has been used Correspondence Analysis (CA), a factor analysis technique that achieves data reduction and noise filters.. The method is meaningful in analyzing cross tabular data in the form of numerical frequencies, and results in an elegant but simple graphical display that permits more rapid interpretation and understanding of the data (Greenacre, 2008). CA has some advantages, such as the symmetric role in the analysis of rows and columns of the data table, and the use of Chi-2 distance properties (Bordons et al. (2004). By means of CA we detected topics (keywords) associated with the years of publication. In this we can find through CA both, emerging research topics and the mainstream. The first are those linked to 2012-2014 and we have considered mainstream those years positioned in the middle of the graph and linked to the whole period.

4- Country specialization:

To determine the specialization of each country in the field of energy efficiency, it was used again Correspondence Analysis. In this case topics (categories WoS) and countries have linked.

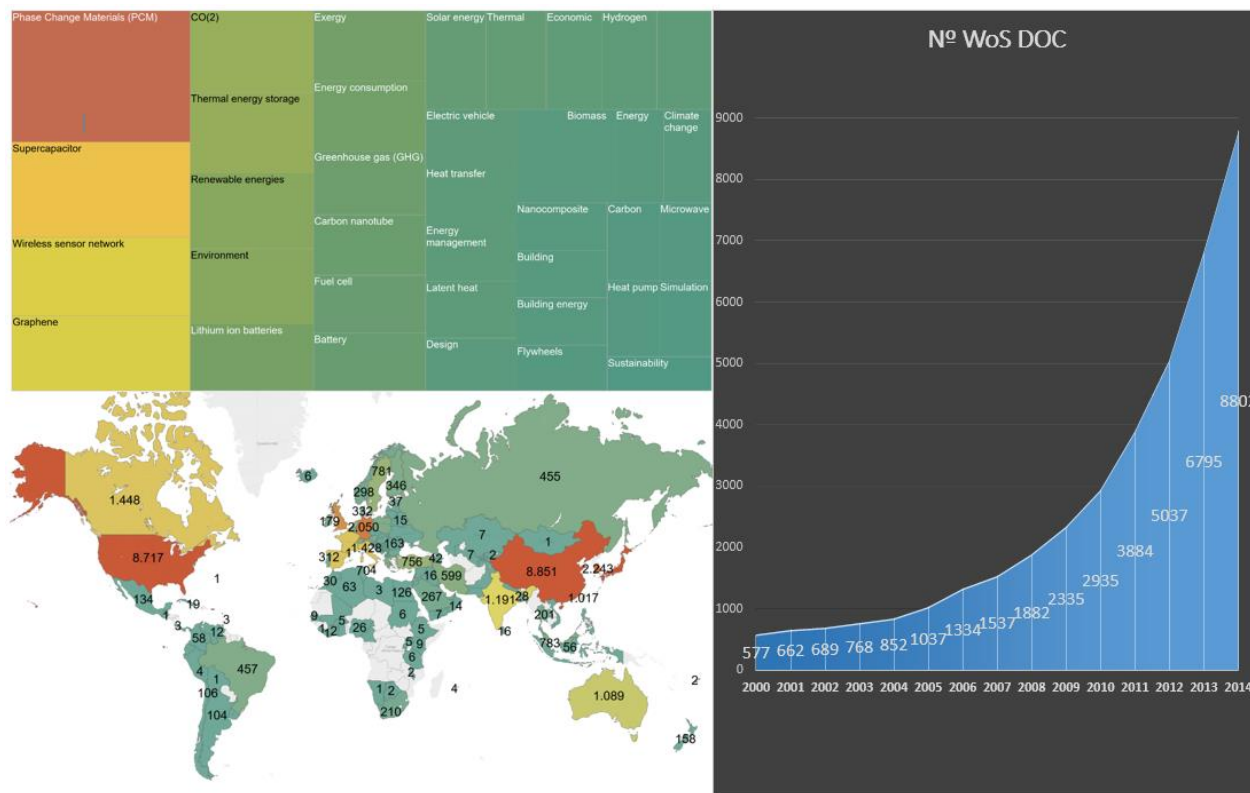
RESULTS

1-Getting bibliometric indicators:

The results of the first phase show that over the period 2000-2014 WoS publications related to Energy Efficiency were 39.126. The number of publications has grown exponentially. The last year has analyzed nearly a quarter of all papers (22.51%). A wide range of countries has generated these documents. USA, China, Japan, South Korea, Germany, England, Spain, Canada, Italy, France, India, Australia and Taiwan are the countries that have higher production in the field

The subject areas (WoS Categories) that have a higher weight are Energy fuels; Electrical electronic engineering; Material science multidisciplinary; Chemical physical; Thermodynamics; Physics applied. The keywords that appear most frequently are Phase change materials (PCM), Supercapacitor, Wireless sensor network, Graphene and CO₂ In Figure 1, the main indicators are presented

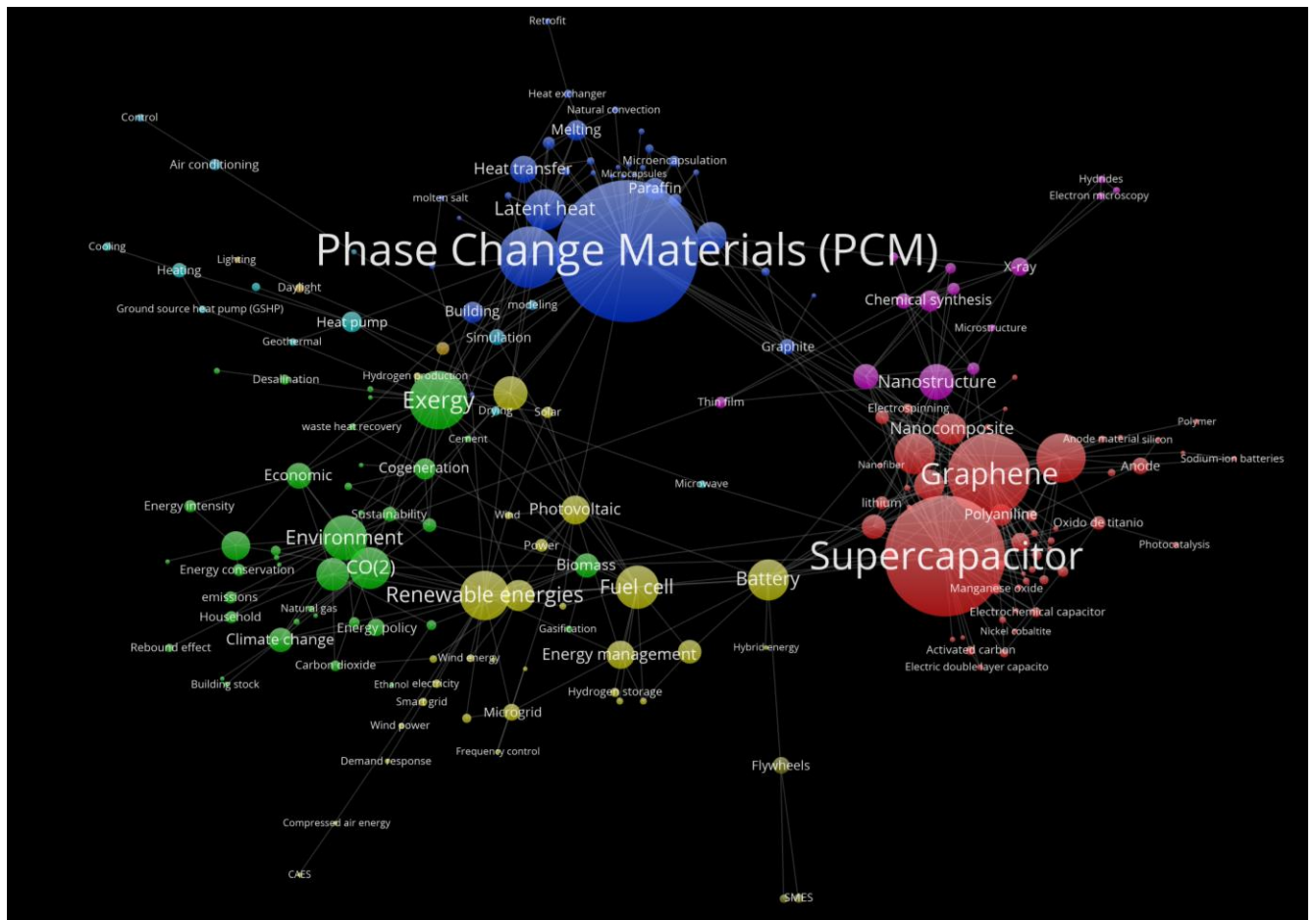
Figure 1. Number of documents per WoS Category, year and country.



2-Revealing the thematic structure of knowledge domain

With regard to the second phase, the co-occurrence analysis of the most frequent keywords has generated a network with 18 clusters (Figure 2). The groups of nodes around Supercapacitor and Graphene (red) are the most prominent. In addition, those nodes that is around PCM (blue) and other groups of nodes like Enviroment (green) or the nodes around the node Renewable energies (yellow).

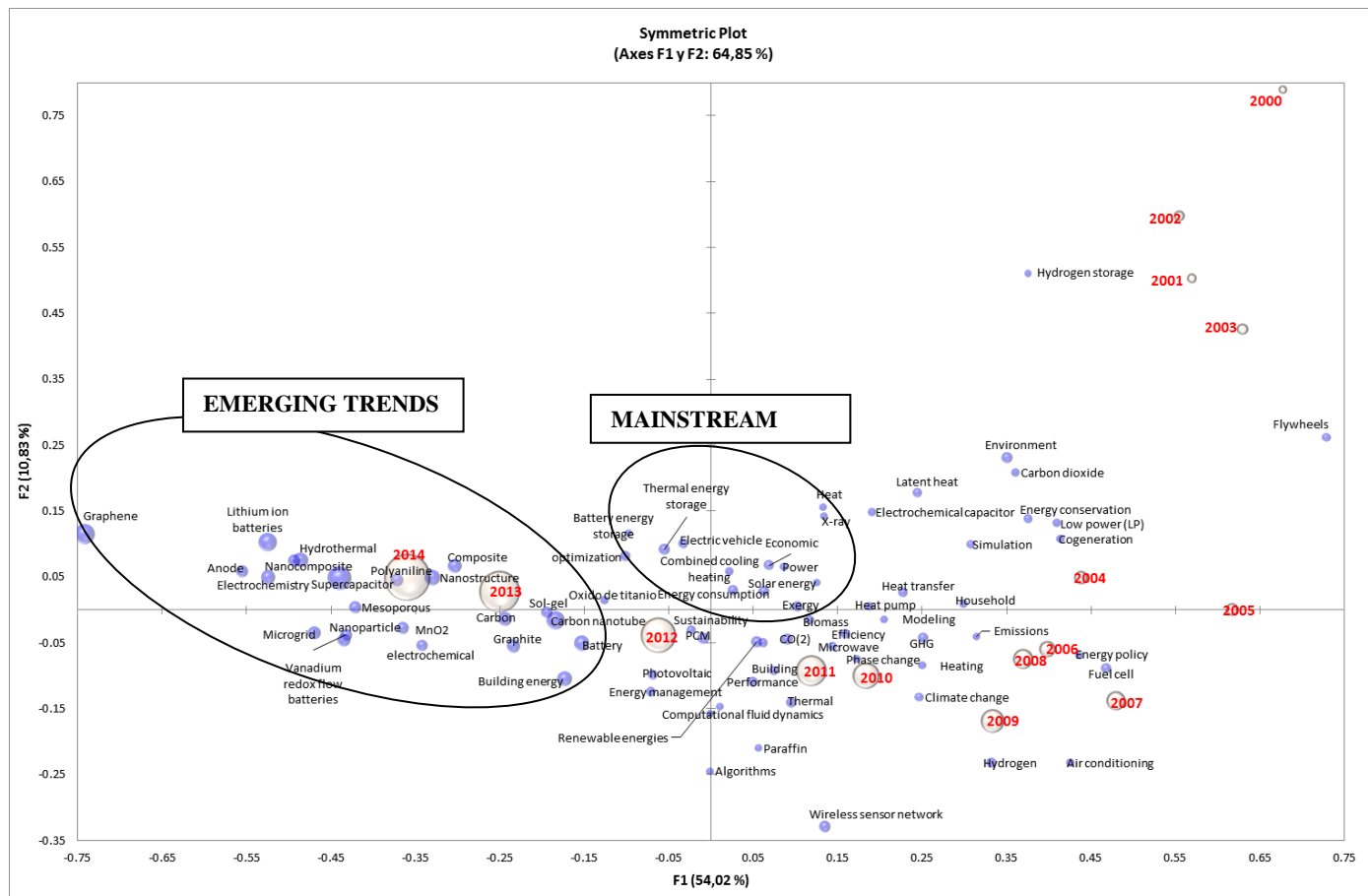
Figure 2. Networks in energy efficiency. Clusterization of topics



3-Identification emerging areas of research:

The figure 3 shows, through an analysis of correspondences, the mainstream and the emerging research areas. The keywords appear link to years of publication. The keywords appear linked to years of publication. White bubbles represent the years and blue bubbles the topics. As can be seen, as Supercapacitor keywords, Grapheno, Lithium Batteries, Graphie, etc., are emerging research fronts of energy efficiency because are related with subject published in the last two years. The topics that appear in the center of the figure (Sustainability, PCM, Solar Energy, CO₂, Combined Cooling Heating, Electric Vehicle) are issues that have received constant attention throughout the period analyzed therefore, we consider these as central or mainstream topics.

Figure 3. Correspondence Analysis: Keywords vs Years.



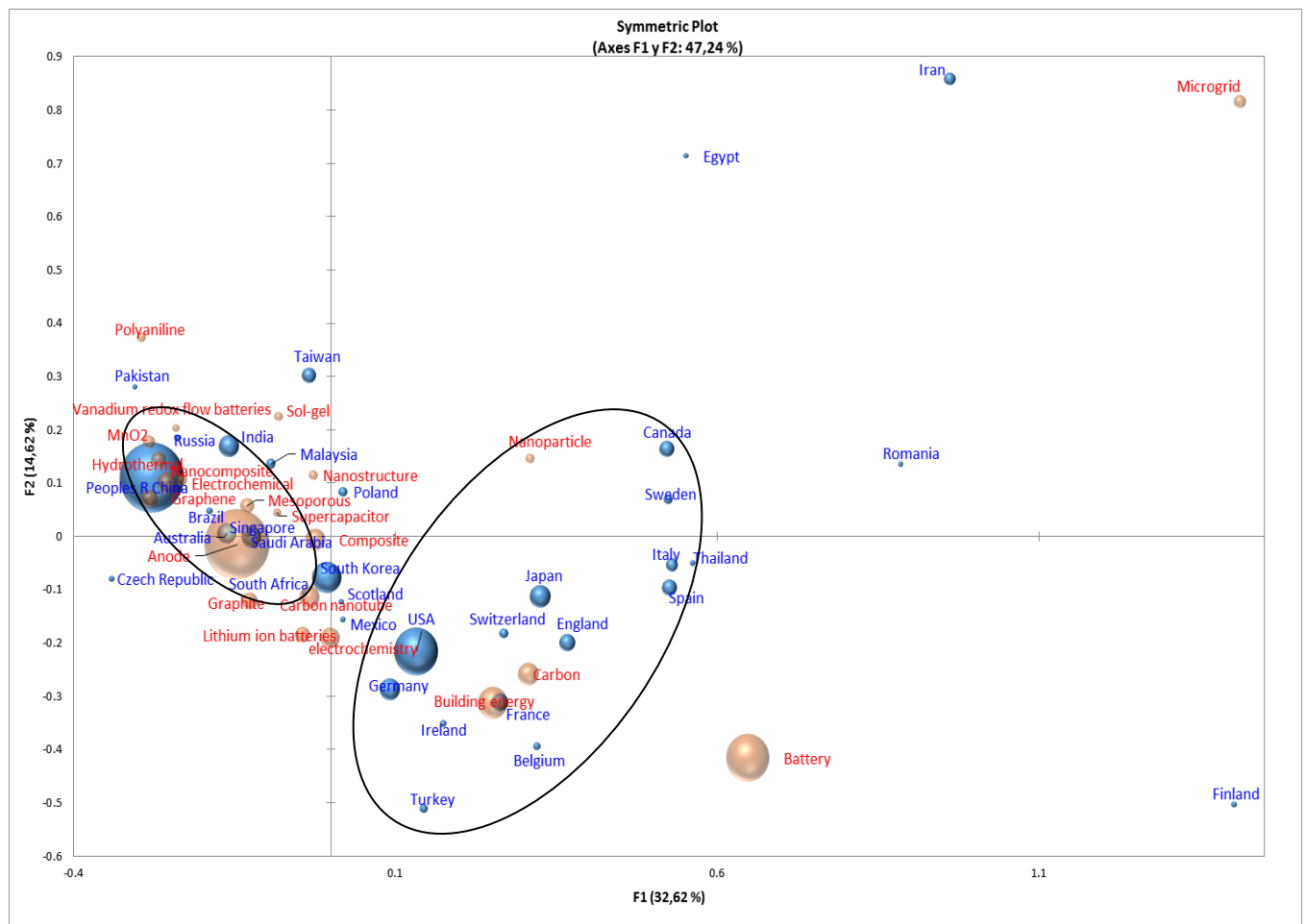
4-Country specialization:

In the last step, we intend to identify the thematic specialization of countries in emerging research areas. The blue bubbles represent countries and orange bubbles the emerging issues. As in Figure 3 the size of the bubble is proportional to the number of papers. The proximity between countries and issues determines the thematic specialization of the country.

The Correspondence analysis has shown that the countries of North America and Europe have similar profiles. Moreover, we have also observed that the BRICS have their own profile. Finally, we have also found that the countries of the periphery has very specific specializations (Iran, Egypt, Romania, Pakistan).

The analysis in this paper has revealed that countries have different thematic profiles when working in emerging research areas

Figure 4. Correspondence analysis by Countries vs Keywords



CONCLUSIONS

The energy efficiency sector is key to the growth and development of the countries. This has favored the growth of research on issues related to energy efficiency that can be traced by analyzing scientific publications. In this regard, our research has shown that scientific production about energy efficiency has suffered an exponential growth over the last fifteen years. In this period the production has risen from 0.04% of publications in Web of Science in 2000, to 0.38% in 2014. Furthermore, the growth of the database in these years was 72% while the field analyzed increase more than 1000%. This shows the growing interest in this subject at the international level.

Considering specialization by the frequency of WoS categories, the most relevant subject are Energy fuels; Electrical electronic engineering and Material science multidisciplinary. However, if the study is performed by keyword frequency, different cluster associated with new materials and energy conservation appear: Phase change materials, Supercapacitor and Graphene (closely linked latter two terms). Other important topics are these focused in Renewable Energy and Environment. It is interesting to detect that, some of these terms as

Graphene or Supercapacitor are emerging topics, because they have appeared in recent years. The volume of documents in these areas is an indicator of current interest in this subject. The other major issues of Energy Efficiency (PCM, Wireless sensor network, CO₂), have been treated consistently over the years.

The last correspondence analysis (country versus keyword) revealed that countries have different thematic profiles when working in emerging research areas. In this respect, they have been detected three types of profiles: a) North America and Europe; b) BRICS and c) countries of the periphery.

REFERENCES

- Bordons, M., Bravo, C., & Barrigon, S. (2004). Time-tracking of the research profile of a drug using bibliometric tools. *Journal of the American Society for Information Science and Technology*, 55(5), 445-461. doi: 10.1002/asi.10397
- Bornmann L, Leydesdorff L (2013) Macro-Indicators of Citation Impacts of Six Prolific Countries: InCites Data and the Statistical Significance of Trends. *PLoS ONE* 8(2): e56768. doi:10.1371/journal.pone.0056768
- Borner, K., Chen, C. M., & Boyack, K. W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology*, 37, 179-255. doi: 10.1002/aris.1440370106
- Calero-Medina, C., & Noyons, E. C. M. (2008). Combining mapping and citation network analysis for a better understanding of the scientific development: The case of the absorptive capacity field. *Journal of Informetrics*, 2(4), 272-279. doi: 10.1016/j.joi.2008.09.005
- Callon, H. M., Courtial, J.-P., & Penan, H. (1995). *Cienciometría: la medición de la actividad científica: de la bibliometría a la vigilancia tecnológica*: Trea.
- Chen, C. M. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57(3), 359-377. doi: 10.1002/asi.20317
- Confraria, H.; Mira Godinho, M. (2015). The impact of African science: a bibliometric analysis. *Scientometrics* (2015) 102:1241–1268
- Dore, J. C., & Ojasoo, T. (2001). How to analyze publication time trends by correspondence factor analysis: Analysis of publications by 48 countries in 18 disciplines over 12 years. *Journal of the American Society for Information Science and Technology*, 52(9), 763-769. doi: 10.1002/asi.1130.abs
- Dore, J. C., Ojasoo, T., Okubo, Y., Durand, T., Dudognon, G., & Miquel, J. F. (1996). Correspondence factor analysis of the publication patterns of 48 countries over the period 1981-1992. *Journal of the American Society for Information Science*, 47(8), 588-602. doi: 10.1002/(sici)1097-4571(199608)47:8<588::aid-asi3>3.0.co;2-p
- Fernández, MT; Sancho, R; Morillo, F; De Filippo, D; Gómez, I. “Indicadores de especialización temática de los países de América latina y el Caribe” En Albornoz y Ratto (eds) *Indicadores de Ciencia y Tecnología en Iberoamérica*. Agenda 2005. RICYT. Buenos Aires. Argentina (2005). pp: 191-209

- Greenacre, M. J. (2008). *La práctica del análisis de correspondencias*. Bilbao: Fundación BBVA.
- Lascurain, M. L., Madera-Jaramillo, M. J., Ortoll, E., & Sanz Casado, E. (2010). Capacidad innovadora de la Comunidad de Madrid a partir de las patentes concedidas entre 1996 y 2007. *Revista Espanola De Documentacion Cientifica*, 33(3), 458-479. doi: 10.3989/redc.2010.3.761
- Mauleon, E., & Bordons, M. (2014). Gender-based indicators for technological activity in Spain based on the study of European patents. *Revista Espanola De Documentacion Cientifica*, 37(2), 13. doi: 10.3989/redc.2014.2.1093
- OECD. (2011). *Towards Green Growth*. Paris.
- Vargas-Quesada, B., & de Moya-Anegón, F. (2007). *Visualizing the structure of Science*. Berlin: Springer-Verlag Berlin Heidelberg.

CHAPTER 4

Societal Impact



Unveiling Research Agendas: a study of the influences on research problem selection among academic researchers

Mariela Bianco^{*}, Judith Sutz^{**}

^{*}mbianbo@gmail.com

CSIC (University Research Council), Universidad de la República, Jackson 1303, Montevideo 11200 (Uruguay)

^{**}jsutz@csic.edu.uy

CSIC (University Research Council), Universidad de la República, Jackson 1303, Montevideo 11200 (Uruguay)

ABSTRACT

Research problem selection is central to the dynamics of scientific knowledge production. Research agendas result from the selection of research problems and the formulation of individual and/or collective academic strategies to address them. But, why researchers study what they study? This paper presents incipient research focused on the way different factors influence the construction of academic research agendas. It takes a researcher-oriented approach relying on opinions and perspectives of a wide range of researchers in all fields of knowledge. The empirical work is carried out in Uruguay, a country in the periphery of mainstream science, whose academic community struggles in search of a balance between the requirements of the world community of scholars and the demands from different national stakeholders. The methodology and research results from this study may be relevant to other countries, at different peripheries. Further, understanding the interplay of influences that shape research agendas is an important tool for policy analysis and planning everywhere.

INTRODUCTION

Research problem selection is central to the dynamics of scientific knowledge production. Research agendas result from the selection of research problems and the formulation of individual and/or collective academic strategies to address them. But, why researchers study what they study? This paper presents incipient research focused on the way different factors influence the construction of academic research agendas. It takes a researcher-oriented approach relying on opinions and perspectives of a wide range of researchers in all fields of knowledge. The empirical work is carried out in Uruguay, a country in the periphery of mainstream science, whose academic community struggles in search of a balance between the requirements of the world community of scholars and the demands from different national stakeholders. The methodology and research results from this study may be relevant to other countries, at different peripheries. Further, understanding the interplay of influences that shape research agendas is an important tool for policy analysis and planning everywhere.

BRIEF BACKGROUND AND SIGNIFICANCE OF A STUDY ON ACADEMIC RESEARCH AGENDAS

Worldwide, transformations in the production of academic knowledge are coupled with greater influences from the extra-academic world (business firms, public policy, civil society and social movements) on research. Nevertheless, changes in the interactions between science and society may lead to contradictory signals when societal pressures over the academic

enterprise clash with internal rules of appraisal of academic quality and productivity. An increasingly evident contradiction is the one between demands for practical relevance of research and dominant universalist forms of academic evaluation (Bianco, et al, forthcoming). “The dominant funding shifts may imply a pressure for more practical relevance, while the rise of performance evaluations has increased the pressure to publish, which may devalue practical concerns and stakeholder interactions” (Hessels, et al, 2011: 555). This contradiction might be stronger in peripheral contexts where science faces many difficulties and research focused on local problems is needed. The search for knowledge solutions for local problems is time consuming, involves multiple partners, and requires the development of circles of trust and understandings among scientists in different knowledge fields. Moreover, publishing in international journals is not the usual outcome of locally relevant research, or peripheral science. For these reasons, the contradiction highlighted by Hessels et al, may be stronger in peripheral contexts.

The selection of a research topic, the formulation of an individual or a collective academic strategy, and the consequent development of different research paths are central academic processes having an impact on which problems are studied and which problems will remain under-researched. Explicit knowledge policies backed by budgetary decisions are one among several influences received by researchers when defining their academic plans: to what extent policies can prevail in the orientation of the scientific enterprise will depend on the relative strength of other influences.

Deciding what to investigate involves choosing a topic among a set of alternatives. This choice results from the interaction of multiple influences and from the consideration of a range of possibilities by researchers and research communities. Factors involved in problem choice are linked to the collective perception about the importance of the issues at stake, the degree of difficulty based on the time and resources required for the inquiry, the relative uncertainty over material returns, the social prestige, and intellectual satisfaction that the choice in question can provide (Ziman, 1987). Among researchers in the natural sciences the selection of the problem may be a function between viability, understood in terms of estimating the time and resources needed for research, and interest, conceived as the expectation of the knowledge contribution (Alon, 2009). However, this simplification conceals a more complex process in which different “sources of influences” and “degrees of influence” interact for the selection of research topics (Cooper, 2009).

There is little evidence-based knowledge accumulated on this subject as previous studies tend to focus on particular fields and/or on specific influences. In applied areas such as agricultural sciences or medical research in which priority setting in agriculture or public health is more necessary and/or justified, there is some background on this topic. A couple of studies dealing with different issues of the scientific profession indirectly address influences on academic research in Mexico (Bensusan et al., 2014) and in selected countries (Teichler et al., 2013). Some analysis of a particular influence -evaluation methods- can be found in a comparative basis between two disciplines (Leysite et al, 2008) or in one particular discipline like demography (van Dalen and Henkens, 2012). Other studies analyzed the influence of researchers’ identities and value orientations on their research agendas (Glenna et al., 2011; Olmos et al., 2013). A recent review highlights what is still not fully understood about the relation between science governance and changes in research contents (Gläser and Laudel, 2016).

THE STUDY

Researchers' choices are influenced by multiple circumstances. Inquiring about the motivations of researchers for their decisions regarding what to study involves taking into account multiple influences. Thus, the questions guiding this study are: What is the relative importance of various groups of factors in the selection of topics for academic research in Uruguay? Are there different researcher profiles based on different research purposes? How are these profiles manifested in different areas of knowledge?

We are studying active researchers in all academic fields totalling about 1300 individuals in Uruguay. After an exploratory phase, an on-line survey will provide the bulk of data organized in different sets: a) demographic and cognitive specialization questions, b) questions on research motivations, c) opinion questions regarding varied criteria for problem choice including institutional mandates, intrinsic academic interests, specific knowledge demands made by stakeholders, incentives from the research funding system and incentives from academic performance evaluation, and d) questions on individual and collective strategies regarding the research topics pursued. Various questions adopt a five-point response scale for the assessment of influences on the selection of research problems. Each influence is an item measured by a numerical scale assuming a continuum from one end meaning "not important" to the opposite meaning "very important;" a separate sixth category is used for "do not know." This format facilitates the comparison of different items in their relative importance among researchers for the assessment of sources and degrees of influence on research.

The goal of the analysis is twofold. The different ways in which individual characteristics are reflected in researchers' choices will be examined with descriptive statistics and multivariate analysis. Relevant dimensions for characterization and comparison are age, gender, stage in the academic career, cognitive area, and institutional affiliation, among others. Further, exploratory factor and cluster analysis will be used to identify patterns of variation of responses and visualize possible groups of researchers regarding their research orientations and sources of influences on research decisions.

The empirical phase of this study has recently begun. Preliminary findings from an exploratory phase show about twenty different criteria suggested by interviewed researchers for the selection of their main research topics. These criteria can be classified according to the following categories:

- a) Cognitive factors mainly derived from the dynamics of knowledge production
- b) Factors associated with STI policies, their instruments and incentives
- c) Demands and interests expressed by non academic actors
- d) Personal characteristics of researchers and career paths
- e) Researchers' expectations regarding the potential use of research results
- f) Individual or collective strategies for career advancement

Surprisingly, exploratory interviews revealed that researchers do not voluntarily speak about the academic evaluation system as an influence on research topic selection. At a first glance, what is academically rewarded or punished by academic performance evaluation is not reported by researchers as a determinant of research direction. This aspect will be inquired with more detail during the survey phase.

We only have findings from exploratory interviews at this time, but most importantly we would like to show the overall approach of this research and potential for replication in similar peripheral contexts. Besides the geographical meaning of the concept “periphery,” we also use a far more general one concerned with the type of problems and approaches that are in the margins of the scientific enterprise and which are better suited to developing contexts. We hypothesize that diverse influences on academic research have a negative impact on contextualized research agendas, threatening, in particular, the propensity of researchers to engage in topics that involve negotiation with non scientists. A robust dataset, as the one resulting from this study, will allow multiple inquiries to remove the veils from research agendas. We believe that ignoring the determinants on research problem selection conceals the fundamental reasons why some research topics remain unstudied, some approaches continue being peripheral and mainstream research gets stronger.

REFERENCES

- Alon, U. (2009). “How To Choose a Good Scientific Problem”. *Molecular Cell* [on line]. Vol. 35, 726-728. Retrieved March 10, 2016 from: <http://www.imbb.forth.gr/people/aeconomou/pdf/HowToChooseGoodProblem.pdf>.
- Bensusán, G., Gras, N., Inclán, D., Rodríguez, E., Valenti, G. and Varela, G. (2014). Reflexiones sobre la evaluación a los investigadores: una mirada desde diferentes perspectivas. Retrieved March 10, 2016 from: http://www.foroconsultivo.org.mx/libros_editados/evaluacion_de_la_evaluacion_subgrupos_individuos.pdf.
- Bianco, M., Gras, N. and Sutz, J. Academic evaluation: Universal instrument? Tool for development? *Minerva* (forthcoming).
- Cooper, J. (2009). Commercialization of the University and Problem Choice by Academic Biological Scientists. *Science, Technology, & Human Values* 34, 629-653.
- Gläser, J. and Laudel, G. (2016). Governing Science. *European Journal of Sociology* 57(01): 117-168.
- Glenna, L.; Welsch, R.; Ervin, D.; Lacy, W. B. and Biscotti, D. (2011). Commercial science, scientists’ values, and university biotechnology research agendas. *Research Policy* 40, 957-968.
- Hessels, L. K.; Grin, J. and Smits, R. E. (2011). The effects of a changing institutional environment on academic research practices: three cases from agricultural science. *Science and Public Policy* 38, 555-568.
- Leysite, L., Enders, J. and De Boer, H. (2008). The Freedom to Set Research Agendas. Illusion and Reality of the Research Units in the Dutch Universities. *Higher Education Policy* 21, 377-391.
- Olmos, J.; Benneworth, P. and Castro-Martínez, R. (2014). “Explaining researchers’ readiness to incorporate external stimuli in their research agendas”. *INGENIO Working Paper Series*

[on line]. Working paper nº 2014-08. Retrieved March 10, 2016 from: <http://www.ingenio.upv.es/sites/default/files/working-paper/2014-08.pdf>.

Teichler, U.; Arimoto, A. and Cummings, W. K. (2013). *The changing academic profession: Major findings of a comparative survey*. Berlin: Springer

van Dalen, H.P and Henkens, K. (2012). Intended and Unintended Consequences a Publish-or-Perish Culture: A Worldwide Survey. *Journal of the American Society for Information Science and Technology* 63, 1282-1293.

Ziman, J. M. (1987). The Problem of "Problem Choice." *Minerva* 25, 95-106.



“If we come out with the wrong answer that may affect investments”: Exploring how evaluators were influenced by political considerations during the assessment of societal impact¹

Samuel, GS. *, Derrick, GE. **

* g.samuel@lancaster.ac.uk
Educational Research, Lancaster University, Lancaster (UK)

** g.derrick@lancaster.ac.uk
Educational Research, Lancaster University, Lancaster (UK)

ABSTRACT

This paper explores the extent to which external, political considerations relating to audit culture behaviour influence panel peer reviewers’ decision-making during the evaluation of societal impact.

INTRODUCTION

Shifts in research policy, and increased squeezing of government spending has meant that research, which was historically viewed as a social endeavour for the public good, is being recast in terms of productivity, ‘economic efficiency’, accountability and delivering ‘value for money’. Such shifts are part of a broader ‘audit culture’ which has emerged in Higher Education Institutions (HEIs) (Power, 1994; Shore & Wright, 2015). A new social contract has now arisen between science and state, which revolves around notions of accountability, relevance and value (Demeritt, 2010); and research evaluation has become a prominent exercise in many industrialised countries (Guená & Martin, 2003). These research evaluation systems, which were once associated with the assessment of scientific output, are now increasingly also associated with the assessment of societal impact as governments endeavour to ensure academic research is both accountable and of benefit to society.

Much research has explored how research evaluation exercises, and the academic environment they produce, have influenced academic and organisational behaviour (Butler, 2003; Harley & Lee, 1997; Henkel, 1999; Leathwood & Read, 2013; Manville, Morgan Jones, et al., 2015; Melo, Sarrico, & Radnor, 2010; Talib, 2002). Less research has explored how such changes in behaviour affect peer reviewer behaviour during evaluation exercises. It is seemingly assumed that even though peer reviewers are also affected by audit culture changes, these evaluators are experts within their own disciplines, having sufficient experience with which to perform an objective, expert-led evaluation that is characteristic of any peer review assessment. However, with the emergence of societal impact assessments, a situation has emerged in which evaluators have no experience in applying this new, untested and unclear criterion (Derrick & Samuel, 2016; Samuel & Derrick, 2015). Without the expertise to conduct an objective, expert-led peer review evaluation, there is a danger that evaluators’ exposure to political pressures and audit culture incentives will influence peer review evaluation. This paper explores the extent to which external, political considerations

¹ This work is supported by the Economic and Social Research Council (ESRC), UK

relating to audit culture behaviour influence panel peer reviewers' decision-making during the evaluation of societal impact.

METHODS

Interviews

Interview participants were purposefully sourced from the UK's Research Excellence Framework (REF2014) Main Panel A (health and medical) evaluation panel.

A total of 215 evaluators were identified and invited to participate in two interviews; one prior to the assessment taking place (pre-evaluation interview); and one after the assessment was completed (post-evaluation interview). The combination of pre- and post-evaluation interviews is key in the utilisation of the in-vitro approach to assessing academic evaluative cultures (Derrick, forthcoming). In total, 62 evaluators agreed to participate in the pre-evaluation interviews (28.8% response rate), and 57 evaluators returned to participate in the post-evaluation interviews.

Interviews were conducted via the telephone, skype, or face-to-face; lasted between thirty-five minutes and two hours; and were recorded and transcribed for analysis. A more in depth structure of the pre-evaluation interview schedule, along with a discussion of the methodology's rigorous qualitative framework has been discussed previously (Samuel & Derrick, 2015). Briefly, interview questions drew on participants' past experience with impact as a prompt to explore their opinions about the importance of evaluating research impact, and its inclusion as a formal criterion in the REF2014. Post-evaluation interviews followed a similar format to those above. Interview themes were based around findings from the pre-evaluation interviews, and explored participants' experiences of the evaluation process.

The codes used in the results below relate to the participant's panel (Main panel=P0; Sub panel 1=P1 and so forth) and their evaluation responsibilities (Outputs and Impact (OutImp); Impact only (Imp); or Output only (Out) control group).

Analysis

An in-depth discussion of the analysis of the pre-evaluation interviews has previously been provided (Samuel & Derrick, 2015). This technique was repeated for the post-evaluation interviews. In brief, analysis used an inductive approach to grounded theory. Such approaches use an exploratory style methodology and 'coding' techniques, to allow concepts, themes and ideas to emerge from the data (Glaser & Strauss, 1967). Duplicate coding by both the first and second author was cross-checked to ensure reliability of data.

RESULTS

Pre-evaluation interviews

The need to demonstrate public value in science

Even before the assessment of societal impact had commenced, and even before the impact case studies had been read and evaluated, evaluators had established an implicit preconceived outcome of the societal impact assessment. This outcome was related to the 'need...to demonstrate the value of science' (P3Out2)(42%). This 'need' served two purposes. One revolved around ideas of 'accountability' (P2OutImp1)(29%). Originating in the realms of the audit culture, and resonating with new notions of a social contract and the reported purpose of the REF2014, for these evaluators, researchers had a 'responsibility' (P4OutImp6) to

communicate research and explain its benefits to society: *'we do really...have an obligation to society'* (P0OutImp1).

The second narrative, which revolved around a *'justification'* (P1OutImp3) for research (35%), was an extension of the above *'accountability'* account in that it was also linked to allocation of research funds. However, beyond merely being answerable to the public for received funding, in this account, evaluators also perceived a need to *'justify'* this received funding. This belief also resonated with narratives of audit culture, accountability and value, as well as with REF2014's aim to *'evidence [the] benefits'* of research (Research Excellence Framework). Similar to the language used by HEFCE themselves², evaluators perceived that *'showcasing'* the value of research conducted within universities during the societal impact assessment provided this very opportunity to evidence these benefits: *'I can see why it is useful for HEFCE to have these cases to show to government that what they have been funding for all these years has had some impact'* (P1OutImp7).

The showcasing of impact, however, went beyond REF2014's stipulated requirement to evidence research benefit. It became an *'advocating mechanism'* to demonstrate the worthiness and value of research receiving public funding over and above other contenders of public money, as well as a tool to demonstrate the validity for further funding as a reward for what had already been achieved. As such embodied the continued political struggle fought by the research community to ensure that science maintains a dominant place in public life and government funding (Kearnes & Wienroth, 2011). Evaluators were aware of the limited resources available to universities and expressed a need to *'argue'* for continued funding: *'resources are getting smaller and smaller - we have to make a reasoned argument as to why we should be funded'* (P1OutImp3). As P4OutImp4 pointed out, *'this is essentially and exercise in convincing the treasury to give us more money'*. In short, the need to convince the government about the benefit of science research was viewed by evaluators as paramount in terms UK science and investment.

Post-evaluation interviews

The political purpose of societal impact

During post-evaluation interviews, evaluators' views about the political purpose of the impact evaluation process had not changed. In fact, similar to pre-evaluation, in the post-evaluation interviews evaluators' narratives about the rationale of assessment still reflected those stipulated by REF2014. Evaluators' continued to articulate the evaluation process as an accountability exercise to evidence research benefit to the government and communicate and educate the public about the emerging societal benefits of research (*'it would be very good...to give lots of good examples of impact of research...the general public are not aware of this. It's a good story, it's a British story and it's worth counting'* (P4OutImp4)).

Evaluators also continued to perceive the impact case studies, and their subsequent evaluation exercise, as an advocacy mechanism to reassure the government that the funding already provided to the research sector had produced societal benefit (*'from the point of view of government...and from maybe people giving money to charities, it's reassuring them...that what they have contributed to is making a difference'* (P1OutImp2)), as well as a need to

² <http://impact.ref.ac.uk/CaseStudies/FAQ.aspx>

influence future government funding allocations (*'we wanted to get money from the treasury for higher education'* (P0OutImp4)).

In the sections below we show how evaluators' uncertainty about how to assess the new, untested societal impact criterion allowed evaluators' perceptions about the societal impact evaluation as an accountability and advocacy exercise to influence the assessment process. These influences to 'showcase impact' manifested in a pattern of generous marking.

Showcasing societal impact through generous marking

Marking generously was a consistent theme discussed by all evaluators in the post-evaluation interviews (*'we scored them high, I thought we were generous'* (P5OutImp1)), including by the control group non-impact (output only) evaluators (*'it is an assumption that everything ought to be three or four star'* (P2Out1)). Therefore, to some degree, the generous marking scheme evident in the peer review assessment of societal impact was unrelated to the newness of the criterion, or evaluators' inexperience in assessing the measure. Rather, it was related, as perceived by evaluators, to a range of both external and internal factors which were influencing both output and impact assessment.

However, separate to the output evaluation, it was the newness of the societal impact criterion, and the lack of clarity regarding how to assess this measure, which helped facilitate the utilisation of the external, political influences of accountability and advocacy as a yardstick for generous assessment in this particular evaluation process. This was because, unlike output evaluation, evaluators were uncertain of how to mark impact, having neither the experience, nor an evaluation precedent, to assess the case studies (*'my expertise to be totally critical about a research paper is much, much more advanced than my expertise to assess whether an impact case study is a really great story or not'* (P5OutImp3)). Evaluators therefore needed a guide for assessment - a benchmark which to follow. Such a benchmark could, and was, drawn from a number of places, for example, from previous personal experience of medical training (*'rule of the hospital, first do no harm – I think that was an approach most of us probably had for the impact'* (P3OutImp2)). Evaluators were also influenced by the external, political factors of accountability and advocacy, which were then also incorporated as benchmarks; evaluators' need to showcase impact became a generous yardstick against which to evaluate impact case studies. This need was further promoted by Main Panel A: *'we were reminded several times that this [the assessment exercise] was about making sure that British research was seen as good'* (P2Imp2).

APPROACHES TO GENEROUS MARKING

The way in which the generous marking scheme was applied provided further evidence that external, political influences related to audit cultures and research evaluation behaviours were behind evaluators' responses to mark generously. In fact, P0OutImp1 explained how the message received from the Main Panel was to showcase impact: *'the major message to me was we had to do a better job of...showing people what we're doing and how important it is'*. Evaluators were encouraged to distinguish their assessment of societal impact ('not assessing in grant-mode') from their usual, more critical, style of marking often used, for example, during grant funding evaluations ('grant mode'). Evaluators were prompted to grade impact highly (all case studies were potentially a four star), unless reason was evident to lower the mark: *'the strategy I came away with was not to think you're in a grant-awarding mood...whatever you have in front of you is good unless proven otherwise'* (P2OutImp2).

Such political influences of advocacy resulted in evaluators becoming *'slightly more liberal in what constitutes evidence'* (P1OutImp2).

This generous approach to assessment was compounded by the marking scale evaluators were required to adhere to: the scale was skewed such that once the threshold of a top mark had been reached, all case studies were marked in this top bracket, whether they were very good, *'impressive'* (P0OutImp5), *'outstandingly impressive'* (P0OutImp5), or *'the best of the bunch'* (P0OutImp2). The purpose of the evaluation thus became a reflection of the external, political influences of accountability and advocacy, rather than of a peer review assessment designed to distinguish between different levels of quality: *'in the end the purpose was to celebrate the excellence of British research, not to discriminate between different types of research and its impact'* (P0P2OutImp1).

Generous marking - making impact assessment 'doable'

The consequence of the generous marking strategy was that the exercise became unable to differentiate between different types of impacts. This was because, as described above, as long as an impact case study could be classified as 'good', it had already reached full marks, and as such, did not require to be distinguished from any other case studies which were at a similar level, or better (*'it wasn't particularly discriminatory because so many people were scored at the top level'* (P1OutImp4)). This approach removed many of the pre-evaluation concerns about having to make value judgments about different impacts (Derrick & Samuel, 2016; Samuel & Derrick, 2015). Thus, in line with a post-evaluation report by RAND, which reported that evaluators perceived the impact component of REF2014 to have *'gone well'* (Manville, Guthrie, et al., 2015), evaluators spoke about the ease of assessment comparative to their original anticipation: *'I think there was a lot of uncertainty...and then when you started doing it, it was surprisingly easy'* (P4OutImp5). The assessment process was therefore announced a success: *'the final report from the REF was almost overwhelming because it was a splendid thing, and it all worked'* (P0OutImp4).

However, this apparent success seemed based on the fact that evaluators' generous marking – a result of, among other factors, the external, political influences of accountability and advocacy affecting the assessment of a new, untested criterion – meant that evaluators did not actually have to assess impact at all. Rather, as described above, marking simply involved weeding out the comparatively poor impacts. Evaluators therefore *'lulled themselves'* into believing it worked (P0OutImp4) when in actual fact, due to the newness of the societal impact criterion, evaluators were politically steered towards an evaluation approach which was externally and politically motivated and less dependent of the opinions of experts that ordinarily provide legitimacy to peer review assessment. In light of this, one evaluator described the exercise as a 'fass': *'the main panel chairman came in and said we'd underscored, so we needed to up scores, which in my opinion made a fass of the whole exercise'* (P6OutImp2).

CONCLUSION

This paper has explored the extent to which panel peer reviewers' decision-making was influenced by the external, political considerations of accountability and advocacy during the evaluation of the new, untested, and unfamiliar criterion of societal impact. It showed that because evaluators did not know how to assess this measure they needed benchmarked parameters within which to evaluate. And because evaluators were not constrained by the

precedents of marking scientific quality, they could draw this benchmark from political motivations outside of academia. With external, political motivations about the need to showcase impact, and to justify the funding of scientific research to policymakers and the public, evaluators' benchmarked parameters became a 'political steer' to err on the side of promotion and generous marking. The peer review assessment of impact was thus affected by external, political motivations influenced by the audit environment itself, rather than being solely dependent on the opinion of expert peers. This has implications in terms of the legitimacy of the peer review process, as well as academic governance.

REFERENCES

- Butler, L. (2003). Modifying Publication Practices in Response to Funding Formulas. *Research Evaluation*, 12(1), 39–46.
- Demeritt, D. (2010). Harnessing science and securing societal impacts from publicly funded research: reflections on UK science policy. *Environment and Planning A*, 42(3), 515-523.
- Derrick, G. E., & Samuel, G. (2016). The societal impact-focused and quality-focused evaluator: Balancing evaluator views about the assessment of the societal impact of health and medical research. *Minerva*, *Accepted*.
- Glaser, B. G., & Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine.
- Guená, A., & Martin, B. (2003). University research evaluation and funding: an international comparison. *Minerva*, 41, 277-304.
- Harley, S., & Lee, F. S. (1997). Research Selectivity, Managerialism, and the academic labor process: The Future of Nonmainstream Economics in U.K. Universities. *Human Relations*, 50(11), 1427-1460.
- Henkel, M. (1999). The modernisation of research evaluation: The case of the UK. *Higher Education*, 38, 105-122.
- Kearnes, M., & Wienroth, M. (2011). Tools of the Trade: UK Research Intermediaries and the Politics of Impacts. *Minerva*, 49, 153-174.
- Leathwood, C., & Read, B. (2013). Research policy and academic performativity: compliance, contestation and complicity. *Studies in Higher Education*, 38(8), 1162-1174. doi: 10.1080/03075079.2013.833025
- Manville, C., Guthrie, S., Henham, M.-L., Garrod, B., Sousa, S., Kirtley, A., . . . Ling, T. (2015). Assessing impact submissions for REF2014: an evaluation. In R. Europe (Ed.). Cambridge.
- Manville, C., Morgan Jones, M., Frearson, M., Castle-Clarke, S., Henman, M.-L., Gunashekar, S., & Grant, J. (2015). Preparing impact submissions for REF 2014: An evaluation. Findings and observations: RAND Europe.
- Melo, A. I., Sarrico, C. S., & Radnor, Z. (2010). The influence of performance management systems on key actors in universities *Public Management Review*, 12(2), 233-254.
- Power, M. (1994). *The Audit Explosion*. London: Demos.
- Research Excellence Framework. REF2014 Impact Case Studies. from <http://impact.ref.ac.uk/CaseStudies/FAQ.aspx>
- Samuel, G., & Derrick, G. E. (2015). Societal impact evaluation: Exploring evaluator perceptions of the characterization of impact under the REF2014. *Research Evaluation*, *Online first*. doi: 10.1093/reseval/rvv007
- Shore, C., & Wright, S. (2015). Audit Culture Revisited: Rankings, Ratings, and the Reassembling of Society. *Current Anthropology*, 56(3), 421-444.

Talib, A. A. (2002). *Performance Measures and Resource Allocation: The Behavioural Consequences of the University Research Assessment Exercise in England*. (Doctor of Philosophy in Industrial and Business Studies), Warwick.



‘Productive interactions’ for societal impact: Developing a research information system for agriculture (RIS-Agric) at Stellenbosch University, South Africa¹

Nelius Boshoff* and Harrie Esterhuysen**

*scb@sun.ac.za

Centre for Research on Evaluation, Science and Technology (CREST) and DST-NRF Centre of Excellence in Scientometrics and Science, Technology and Innovation Policy (SciSTIP), Stellenbosch University, Private Bag X1, Matieland, Stellenbosch, 7602 (South Africa)

** hesterhuysen@sun.ac.za

Centre for Research on Evaluation, Science and Technology (CREST) and DST-NRF Centre of Excellence in Scientometrics and Science, Technology and Innovation Policy (SciSTIP), Stellenbosch University, Private Bag X1, Matieland, Stellenbosch, 7602 (South Africa)

ABSTRACT

Assessments of the societal impact of research rely on primary data as input to the relevant indicators. In the current project, the data collection for indicator construction is guided by the SIAMPI approach to societal impact assessment. According to SIAMPI, the achievement of societal impact involves two elements: stakeholders and productive interactions. The latter includes three kinds of interactions: direct personal interactions; indirect interactions where contact is mediated by a material or human ‘carrier’; and financial interactions that refer to the economic exchanges between researchers and stakeholders. The question to be explored is whether the assessment of the societal impact of research can be facilitated by capturing any direct, indirect and financial interactions between researchers and societal stakeholders in a research information system. CREST/SciSTIP is developing such a system, called RIS-Agric, in cooperation with the Faculty of AgriSciences at Stellenbosch University and four agricultural research funders in South Africa that represent different agricultural commodities. A RIS with productive interactions at its core – and the funded research project as the unit of observation – will produce the necessary data for indicators linked to the societal impact of research, in addition to generating insights about the ‘pathway’ from research to societal impact. The indicators need to be ‘socially robust’ in the sense that their validation processes should include the research beneficiaries and their representatives. In doing so, RIS-Agric will produce a useful, standardised measurement mechanism that is aligned across the different commodities. The focus on productive interactions will be supplemented by other methods in order to enable the quantification of societal impact in the form of ‘impact scores’.

BACKGROUND

Organisations that fund research and those that perform research are subject to the same accountability pressure: to provide evidence of value-for-money, often expressed as impact, in relation to the research investment made. Current methods to assess the scientific impact and contribution of research, although not perfect, are well-established. It includes counting the scientific outputs of research and tracking the numbers of citations to publications in the

¹ This work is supported by a grant from the DST-NRF Centre of Excellence in Scientometrics and Science, Technology and Innovation Policy (SciSTIP), South Africa.

scientific literature. Such methods, associated with alluring descriptors like ‘journal impact factors’ and ‘normalised citation rates’, gained prominence largely because of the existence of two citation databases (Web of Science by Thomson Reuters and Scopus by Elsevier). The latter serve as readily-available sources for calculating the necessary indicators of scientific impact based on secondary data analysis, without any need for primary data collection.

The assessment of the societal impact of research, on the other hand, is still in its infant shoes and very much reliant on primary data collection. Some attempts have been made to explore the relevance of available databases in the domain of social media (Plum Analytics, Altmetrics.com and ImpactStory) for purposes of research impact. However, given the weak but positive correlations found between citations and so-called ‘altmetric’ scores, it is concluded that although social media statistics do measure some form of impact it is unclear as to exactly what kind of impact is indicated (Costas, Zahedi & Wouters, 2015).

Assessments of societal impact, therefore, cannot escape the need for primary data collection to construct the relevant indicators. Initiatives are required to systematically collect, organise, store and critically reflect on the data that are needed. In the current study, the data collection for indicator construction is firstly guided by the SIAMPI method for societal impact assessment (Spaapen & Van Drooge, 2011). This approach is process-orientated and not aimed at quantification or providing an overall judgement. It is a learning tool to understand how and under what conditions certain interactions between the research and the research stakeholders become ‘productive’, i.e. whether *any* changes are effected in stakeholders as a result of the research, and why and how. The focus on productive interactions is supplemented by two quantitatively oriented approaches, namely impact oriented monitoring (Guinea et al., 2015) and social return on investment (Social Value UK, 2012), as they provide means to assign ‘values’ to societal impact. Quantification – although contestable – cannot be avoided under the present-day paradigm of ‘new public management’ where the governance of research at universities and funding organisations largely occurs through the management of numbers (Elzinga, 2012).

The above three approaches for assessing the societal impact of research feature in a localised research information system for agricultural research (RIS-Agric), which CREST/SciSTIP is developing in cooperation with three agricultural research funders in South Africa. The selected funders represent different agricultural commodities (e.g. the wine industry and deciduous fruit industry). RIS-Agric is currently in its planning phase. The objective of this paper is to communicate the initial conceptualisation of the idea.

The three approaches to be used are discussed next.

Approach 1: SIAMPI and the tracking of productive interactions

The SIAMPI approach (*Social Impact Assessment Methods* for research and funding instruments through the study of *Productive Interactions* between science and society) was developed by Jack Spaapen and his colleagues as part of the Seventh Framework Programme of the European Union (EU) (Spaapen & Van Drooge, 2011). According to the approach, the achievement of societal impact involves two main elements: stakeholders and productive interactions. The latter is broken down into three kinds of interactions. First are direct personal connections, which involve face-to-face contact or interactions over the phone, email or video-conferencing. These include meetings, conferences and chance encounters. The second kind of interaction involves indirect encounters where contact is mediated by a material or human ‘carrier’. Examples include articles, reports, guidelines, codes of practices

and other individuals who act as intermediaries. In the last instance we find financial interactions, which refer to the economic exchanges between researchers and stakeholders. These typically take the form of research contracts or financial contributions. Stakeholders, the other main element in the approach, are broadly interpreted to include “anyone who takes part in the iterative process that induces the results of research into social impact” (SIAMPI, 2011, p.6).

Productive interactions are considered a necessary condition for societal impact; in other words, societal impact cannot occur without at least one of the three kinds of interactions between researchers and stakeholders also being present. According to the SIAMPI approach, an awareness of productive interactions is needed in order to conduct an assessment of societal impact. The impact assessment basically tracks the productive interactions or, more appropriately, any relevant interactions between the researchers and stakeholders in order to determine which of these can be deemed productive. Productive interactions are those that induce instances of behavioural change, uptake and use in the stakeholder domain. The SIAMPI approach thus prioritises the processes that create impact and not the impact itself. By doing so, the attribution problem in research impact assessment is side-stepped and the contribution aspect of research emphasized instead (SIAMPI, 2011).

In SIAMPI, information on interactions is primarily collected through the analysis of project documents and by interviewing researchers. Information on the changes effected by the three categories of interactions, together with information on the stakeholders and their conditions, can be obtained through individual interviewing and focus group discussions.

Approach 2: Impact-oriented monitoring of research projects

Impact-oriented monitoring (IOM) is considered “a novel methodology for identifying and assessing the impacts” of funded research projects (Guinea et al., 2015, p.1). The methodology was originally developed for EU-funded research projects in international public health but may also apply to other research fields like agriculture. The methodology is aimed at the development of a project-based research information and assessment system to support a number of operations typically required of research funders: monitoring of project performance; identification of potential and real impacts; assessment of impact; and generation of relevant data for further *ex-post* evaluations. The methodology involves the collection of structured information during and after the project life, based on a theoretical framework. Both the producers and users of research (i.e. project coordinators or grant-holders and the intended beneficiaries) provide information.

IOM has two interrelated components: the theoretical framework and an impact monitoring system. Each, in turn, is divided into separate elements. The theoretical framework, which guides the collection of data, is composed of a conceptual model that follows a logic model approach, and a set of impact categories informed by the Payback Framework (Donovan & Hanney, 2011). Three data collection tools, together with an assessment tool, comprise the impact monitoring system:

- a project results framework which is a simplified (and ‘living’) version of the logic model and updated by the grant-holder at set intervals (at project start and mid-term and final reporting);
- a grant-holder survey also to be completed at three intervals (at mid-term and final reporting and three years after project); and
- an end-users’ opinion survey to be completed by the intended beneficiaries at the end of the project.

The assessment tool is a scoring matrix which assigns an overall score to each impact category of the payback framework, thereby allowing for a comparison both within and between projects.

Approach 3: Social return on investment

“Social return on investment (SROI) is the application of a set of principles within a framework that brings about consistency” in how this value is measured (Social Value UK, 2012, p. 7). SROI measures changes (social, environmental and economic outcomes) that occur as a result of activities. The primary aim of an SROI analysis is to determine how much value (positive or negative) has been created or destroyed for those that are affected or changed by an activity. SROI represents this value as a monetary based ratio.

The approach follows a number of stages of analysis: establishing the scope and identifying stakeholders; mapping outcomes; evidencing outcomes and attaching value; establishing impact; calculating the SROI; and “wrapping up” (reporting, using and embedding outcomes) (Social Value UK, 2012).

Once the scope of the evaluation has been defined and the stakeholders identified, the outcomes are mapped. Outcomes reflect the criteria that a project would like to achieve and are initially detailed either in a theory of change or as part of a logic model (impact plan). Outcomes are present when change occurs. Change is thus captured through indicators. When identifying indicators stakeholders should be consulted first as they will most likely be able to explain how they know that change has happened to them (or not) (Social Value UK 2012, p. 38).

However in SROI it is not enough to just measure an outcome. It is also important to determine its financial value. As with prices in normal day life the value of an outcome is a proxy (Social Value UK 2012:45). Once the outcome values have been determined the deadweight and attribution (to others) are estimated and deducted. Deadweight is the amount of outcome that would have happened even without the evaluated/projected activity taking place. The overall value of impact is then calculated, after which future value is projected through the subtraction of drop-off. This is followed by the calculation of net present value (discounting), which is then finally followed by a calculation of the SROI ratio (present value divided by value of inputs) (Social Value UK, 2012).

SROI is claimed to be much more than just the ratio it produces. It is a ‘story of change’, that includes case studies and qualitative, quantitative and financial information, on which decisions can be based (Social Value UK, 2012).

Towards the development of RIS-Agric

RIS-Agric will incorporate all three above approaches to research impact assessment, as is illustrated in Figure 1 below. Although the approaches are portrayed as three separate activities, their respective elements, information requirements and data collection processes are not entirely independent but can be aligned for greater synergy and cost-saving.

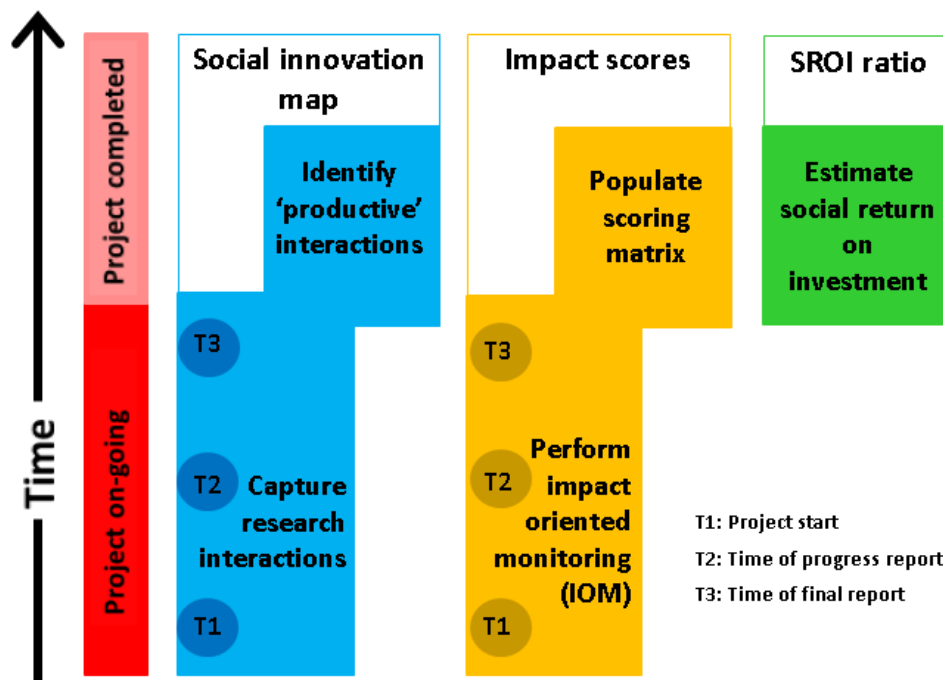


Figure 1. Working structure of RIS-Agric.

From a research perspective, a key question to be addressed through developing RIS-Agric is whether the assessment of the societal impact of research can be facilitated by capturing any direct, indirect and financial interactions between researchers and societal stakeholders in such a research information system (RIS). A RIS with productive interactions at its core (the blue part in Figure 1) – and with a funded research project as the unit of observation – will produce data to construct what can best be called a map of social innovation and produce insights about the 'pathway' from research to societal impact. The data input for generating a social innovation map may be as simple as a dataset consisting of two columns only – the identified interactions and their associated effects.

The relationship between indicators and indications produced during the 'project on-going' stage, on the one hand, and those produced during the 'project completed' stage, on the other hand, also requires systematic scrutiny. Can the first eventually serve as a proxy for the second, thereby reducing the cost of assessment? Equally important is the degree of correspondence between the two sets of overall values (impact scores and SORI ratio) and the social innovation map, all to be produced after project completion.

However, the purpose of RIS-Agric is not only to satisfy research curiosity and enhancing research assessment practice, but also to address needs of funders. Because data pertaining to societal impact or the pathway to such impact will be collected during set stages (T1, T2 & T3) of the project life, funders will have access to real-time data for project monitoring and external communication. Given that most agricultural research funders in South Africa receive voluntary levies from agricultural producers, they are constantly challenged to defend their integrity by showing value for money. Funders often support a portfolio of projects that are 'not up for the vote' in order to sustain their industry's market access. A useful assessment tool for them would be a matrix which shows the importance of research that might not have an immediate impact or that only has a mitigating impact. These expectations are addressed in the working structure of RIS-Agric.

RIS-Agric also seems to bear some resemblance with Researchfish, the UK-based “service for the reporting of outcomes to enable research impact tracking” (<http://www.researchfish.com>), and which is used by the UK research councils and medical charities. In Researchfish, the outputs, outcomes and impact of research – all of which can be attributed to individual grants – are captured by grant-holders (researchers) onto an online platform. This coincides with what RIS-Agric aims to achieve with two of its activities, ‘capture research interactions’ and ‘perform impact oriented monitoring’ (Figure 1). Researchfish uses a broad set of 16 questions consisting of 175 sub-questions for that purpose. In a recent evaluation of Researchfish (Hinrichs, Montague & Grant, 2015) it is stated that the data supports a number of aims: *advocacy* and making the case for research; *accountability* to the funding sources; *analysis* of ‘what works’ in research funding and creates impact; and *allocation* of future funding. All of these equally apply to RIS-Agric.

However (as far as could be established), in Researchfish, the stakeholder communities and intended project beneficiaries do not generate any information on the outcomes and impacts as specified by the researchers. In the case of RIS-Agric, all indicators and indications to be generated also need to be ‘socially robust’ in the sense that their validation processes should include the research beneficiaries or their representatives (Klenk & Hickey, 2013).

In summary, through the development of RIS-Agric which is still in an infant stage, a system is envisaged which will support agricultural R&D management decisions through a useful, standardised measurement mechanism that is aligned across different agricultural commodities. This will enable the funders to measure and report on the contribution of their research to societal impact, and use each other as a benchmark. Valuable lessons are also expected to be learned through engaging with a documentation system for societal impact that is currently being developed by Wolf et al. (2014) for organic agriculture in Germany.

REFERENCES

- Costas, R., Zahedi, Z., & Wouters, P. (2015). Do “altmetrics” correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology*, 66(10), 2003–2019.
- Donovan, C. & Hanney, S. (2011). The ‘Payback Framework’ explained. *Research Evaluation*, 20(3), 181–183.
- Elzinga, A. (2012). Features of the current science policy regime: Viewed in historical perspective. *Science and Public Policy*, 39, 416–428
- Guinea, J., Sela, E., Gómez-Núñez, A.J., Mangwende, T., Ambali, A., Ngum, N., Jaramillo, H., Gallego, J.M., Patiño, A., Latorre, C., Srivaniachakorn, S., & Thepthien, B. (2015). Impact oriented monitoring: A new methodology for monitoring and evaluation of international public health research projects. *Research Evaluation*, 24(2), 131–145.
- Hinrichs, S., Montague, E. & Grant, J. (2015). *Researchfish: A forward look. Challenges and opportunities for using Researchfish to support research assessment*. Research report 2015/11 produced by the Policy Institute at King’s College London.
- Klenk, N.L., & Hickey, G.M. (2013). How can formal research networks produce more socially robust forest science? *Forest Policy and Economics*, 37, 44–56.

SIAMPI. (2011). *Social Impact Assessment Methods for Research and Funding Instruments through the Study of Productive Interactions*. Retrieved from <http://www.siampi.eu/> (retrieval 14 July 2016).

Spaapen, J., & Van Drooge, L. (2011). Introducing ‘productive interactions’ in social impact assessment. *Research Evaluation*, 20(3), 211–218.

Social Value UK. (2012). *A guide to social return on investment*. Retrieved from <http://socialvalueuk.org/what-is-sroi/the-sroi-guide>

Wolf, B., Szerencsits, M., Gaus, H., Müller, C.E., & Heß, J. (2014). Developing a documentation system for evaluating the societal impact of science. *Procedia Computer Science*, 33, 289–296.



Publication patterns in UK research assessment 1992-2013¹

Jonathan Adams*, Tamar Loach* and Martin Szomszor*

[*j.adams@digital-science.com](mailto:j.adams@digital-science.com), t.loach@digital-science.com, m.szomszor@digital-science.com

Digital Science, 4 Crinan St., LONDON N1 9XW, UK

SUMMARY

For 30 years, outputs published by UK researchers have been submitted for cyclical assessment by peer panels, creating a unique longitudinal dataset. This is the first analysis covering 921,254 submitted outputs and 36,244 case study references across 25 years, five assessment cycles and both academic impact and economic/societal impact. For submitted outputs, document types shift towards journal articles across time. The time-spread of outputs is skewed to the most recent publication years in early RAE cycles, a pattern not reported at the time but one that then changes synchronously for science and engineering but not for social science or humanities. The skew and later changes are cohesive across disciplines and institutions. For impact case study references, the time-spread of the earliest of the six references for each study is uniform for science and engineering, although the overall time-spread for references is skewed. About 42% of case study references with DOIs can be identified as RAE/REF submitted outputs, at about the same rate in every publication year for 1996-2013. Implications for assessment are discussed.

INTRODUCTION

How do researchers select publications for assessment, as evidence of their achievement and impact? This study refers to two overlapping sets of publication data, each presented for distinct assessment purposes. The objective of this report is to analyse submission and referencing patterns to arrive at a better understanding of their selective use. The comparative data sets are defined as:

- **Submitted outputs:** outputs (primarily printed publications) submitted in assessment cycles as evidence of academic excellence. The RAE/REF requires four outputs be submitted for each researcher.
- **References:** outputs included in REF2014 as impact case study references. REF2014 differed from the prior RAE format in adding four-page case studies of the societal and economic impact of research, including six references to the underpinning research.

¹ This work was supported by the Higher Education Funding Council for England

Four questions are addressed:

- a. What is the relationship between submitted outputs (evidence of academic impact) and case study references (evidence of societal and economic impact), noting that academic impact can be conceptually associated with research innovation? (Breschi et al., 2006)
- b. What do the data say about the selection of outputs for assessment?
- c. How did selection behaviour respond to assessment?
- d. How homogeneous was that response, by discipline and institution?

METHODS

The UK Research Assessment Exercise (RAE) started in 1986 and settled into a standard template from RAE1992 (UFC, 1992). A peer panel reviewed a portfolio of evidence covering a short census period for a group of researchers in subject categories called Units of Assessment (UOAs). That portfolio included data on outputs (books, journal articles, conference papers, software, art, videos and other).

Data for analysis were sourced from the RAE archive sites maintained by the Higher Education Funding Council for England (HEFCE). The dataset is shaped by the varying length of RAE census periods (the acceptable period for publication cover dates) and by improving technology (e.g. electronic submission). For example, the RAE1992 census period ran from January 1988 for arts and humanities, January 1989 for other subjects and added a part-year with a closing date of June 30, 1992. Census periods for later cycles are mutually exclusive full calendar years.

Each assessment cycle received data for some 50,000 UK-based researchers, across 150 HEIs and a gradually reducing number of between 72 and 36 UOAs. The RA2/REF2 output section of each assessment database contains about 200,000 records of outputs submitted as evidence of research achievement at the time. The combined data set of 921,254 outputs selected for assessment is a unique longitudinal perspective on national research activity.

The Research Excellence Framework (REF2014) followed the RAE but added four-page case studies of socio-economic impact, each supported by reference publications. The census period (1993-2013) overlapped with prior RAE cycles. There are 6,737 non-confidential impact case studies (of 6,975 original documents) available for analysis. The six output references per case study were extracted and normalised by Digital Science, and Thomson Reuters processed items for matches against *Web of Science*TM producing a database of 36,244 identifiable case study journal article references with dates that overlapped assessment cycles from RAE1996.

Submitted outputs and impact case study references were paired to enable an analysis of their overlap. For journal articles, and some other outputs, this was done by matching publication records in both sources to unique DOIs to compare automatically. Normalised REF case study reference records were compared to CrossRef and matched with DOIs where these were missing. Matching publication records to DOIs was also applied to early RA2 submitted outputs where no DOI had been included in original records. The DOI system started around 2000, so there is a deficit of DOIs for earlier publications and not all outputs are journal articles with discoverable DOIs. To establish the overlap between the sets of prior submitted outputs and the set of impact references, the augmented RA2/REF2 submission records were searched for the DOIs of those case study references to which a DOI could be matched. Case

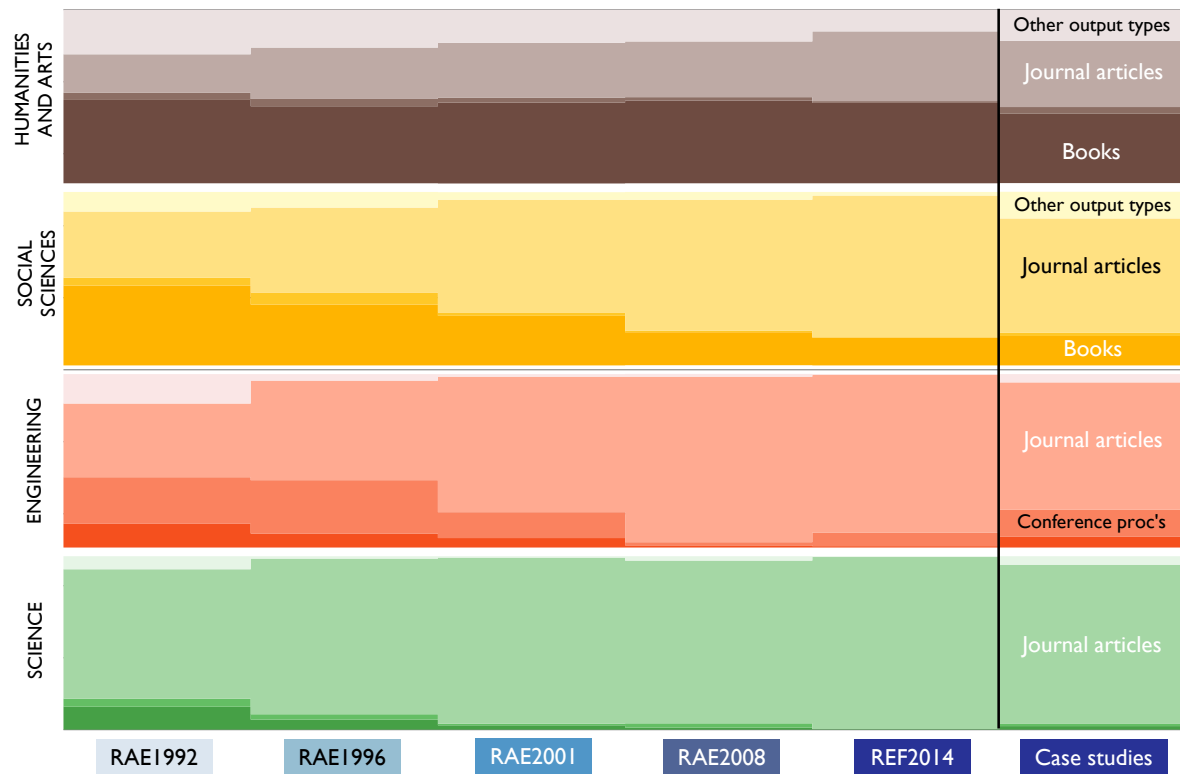
study references without DOIs were processed manually and document-type frequency was summarised for comparison with submitted outputs.

Data were initially processed at UOA level (the disciplinary structure of the RAE/REF) and then a set of higher-level categories was used to overcome the change in UOA count in successive cycles. Data were aggregated into four domains driven by similarity in publication usage (biomedical and physical sciences; engineering and technological sciences; social and economic sciences; and humanities, and visual and performing arts) which differ slightly from the REF Main Panel structure, which combines physical sciences with engineering (REF Panel B), leaving biomedical and clinical sciences (Panel A). For HEI analyses, data were aggregated to the set of HEIs that currently exists. The number and structure of HEIs has changed over the period of analysis, with new foundations, some mergers (e.g. Manchester and UMIST) and some splits (e.g. University of London).

RESULTS

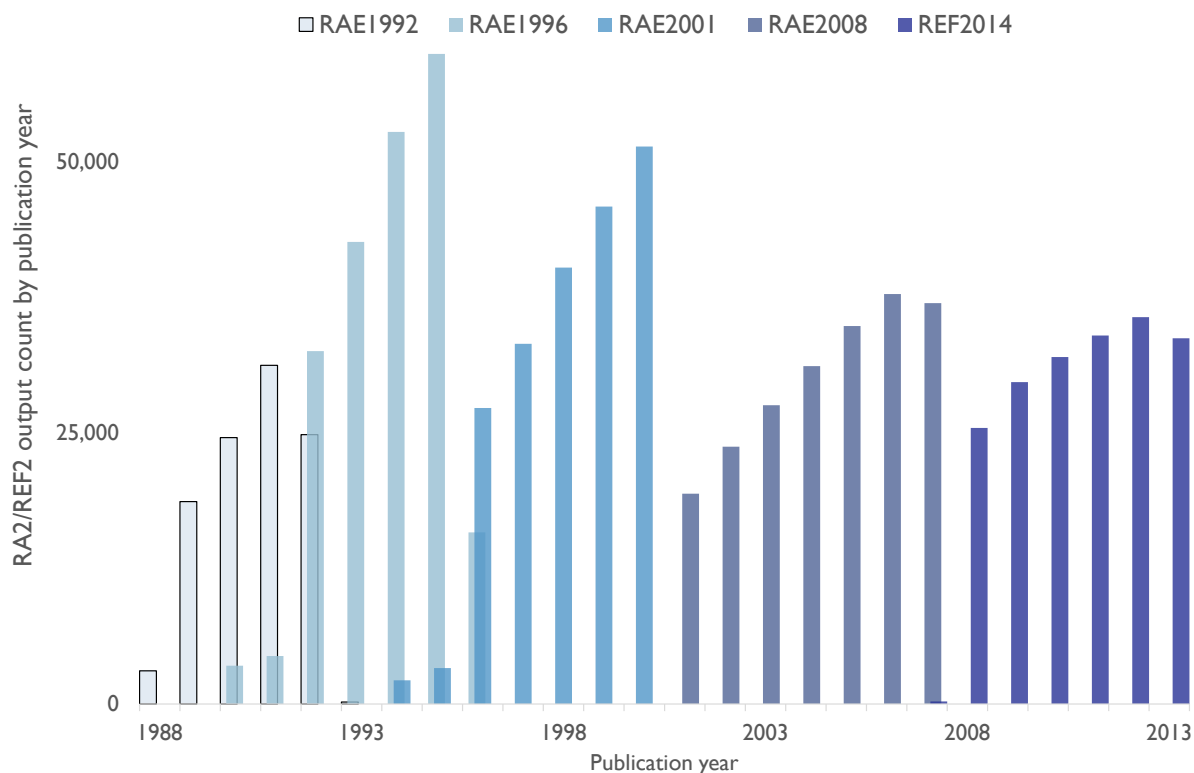
The balance of submitted output types changes over time. There is a changing diversity of output types at the level of science, engineering, social science and humanities. For 1992 and 1996, a preference (in the sense of making a relatively frequent selection) existed for journal articles among scientists; engineers preferred proceedings; the social scientists and scholars in the humanities preferred the monograph; visual and performing arts used specialised media. As early as RAE1996, engineers' preference shifted from conference proceedings and social scientists shifted from books, towards journal articles. In 2008 and 2014, journal articles became the predominant submitted output in all areas except humanities. In counterpoint, the diversity of output types in the REF2014 case study references is more diverse than the REF2014 submitted outputs. It includes a greater proportion of monographs, conference proceedings and non-print outputs and resembles the balance of output types seen in RAE1996 more closely than REF2014. (Figure 1)

Figure 1. Output types as a share of total outputs submitted as evidence of quality in five successive UK research assessment cycles. Subject groups are clusters of a variable number of UOAs in different cycles.



Submitted output publication dates are skewed. The publication dates of submitted outputs are skewed towards the more recent years in each cycle. RAE1992 data reveal a marked time-skew of submitted outputs towards the most recent publication dates for that cycle (1992 was a ‘half-year’ with a census cut-off at 30 June). This skew persisted but was gradually modified in later cycles. Note that a change in the number of years in each cycle (from a core of four years to five to seven to six) affects the height of the curve. The last full year of the census period provided the greatest number of publications submitted for assessment in RAEs 1992, 1996 and 2001. In RAE2008, however, the penultimate year (2006) is the peak publication point for submitted outputs. In REF2014, the time profile is more evenly distributed and 2011 and 2012 publication volumes both exceeded 2013. (Figure 2)

Figure 2. For each UK research assessment cycle, the count of submitted outputs published in each year of the relevant census period (see text for notes on varying census periods).



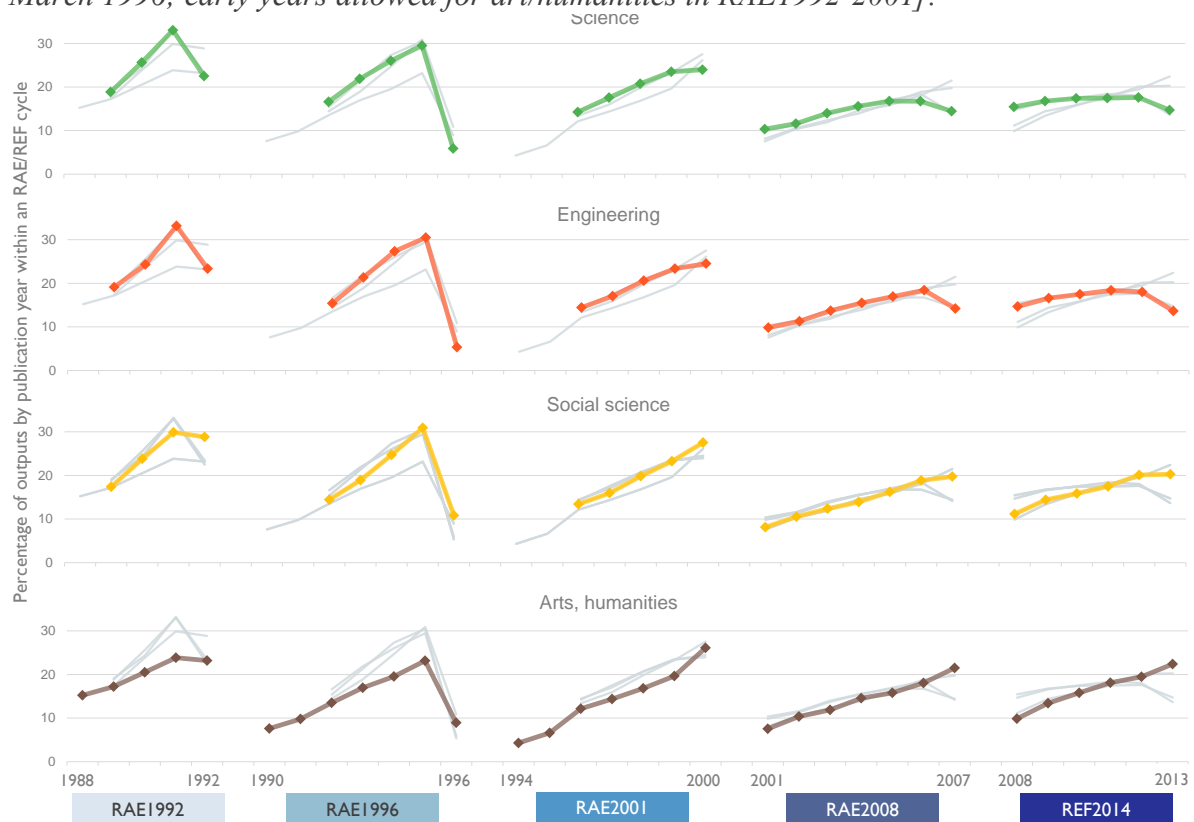
There is no evidence that a shortage of material constrained the selection of outputs. First, HEIs reported that submitted outputs for RAE1992 were drawn from a total pool of 787,138 potentially eligible outputs. Second, Thomson Reuters *Web of Science* database records 90,000 UK-authored journal articles per year so the sum across each census period would exceed requirements. Third, non-indexed journals, conference proceedings, books and other outputs add to the indexed article total. Thus, the selection appears to be deliberate choice, initially skewed to recent outputs and yet subject to successive modification.

Behaviour in selecting submitted outputs is consistent across UOAs and HEIs. The net values for the database in each cycle aggregate many independent submission choices. The preferences may be homogeneous and generic or may be the smoothed outcome of combining varying cultures and missions. To explore this across UOAs and HEIs, two comparative analyses were created by calculating the percentage of total outputs by publication year for each UOA in each cycle and for each of the 150+ submitting HEIs. The median value in each year and the upper and lower quartiles bounding that median indicate the spread across units (UOAs and HEIs). This analysis showed that the 1989-2013 profiles for UOAs and HEIs are similar both to one another and to the overall picture. There is remarkably low variance, with quartiles around the median that are relatively tightly bounded compared to the year-to-year change in medians, and with a narrowing inter-quartile spread in later cycles. (*Published as supplementary material in report to HEFCE, 2016*)

There is a cultural divide in the selection of submitted outputs. Initial cultural distinctions in output selection for assessment are displaced by convergence on journal articles (Figure 1). The analyses underpinning Figure 2 provide other statistical information on variance. By year

of publication, the dataset average is lower than median values for individual UOAs and HEIs: this disparity implies UOA-related differences that then affect HEIs according to their portfolio. To explore this, the data were aggregated into broad subject groups as in Figure 1. At group level, profiles are similar in RAE1992; in RAE2001, a terminal inflection appears for science and engineering; and by RAE2008 a divergence between ‘arts’ and ‘sciences’ becomes clear. Finally, in REF2014, science and engineering peak across two penultimate years and dip in the final census year while social sciences and humanities/arts retain the skew of earlier cycles. (Figure 3)

Figure 3. *The percentage of outputs submitted in each UK research assessment cycle that were published in each relevant census year, disaggregated by major subject groups (as Figure 1). [Curve heights are affected by variable year count in each cycle (see Figure 2); terminal drop in RAE1992, RAE1996 due to publication cut-off at 30 June 1992 and 31 March 1996; early years allowed for art/humanities in RAE1992-2001].*

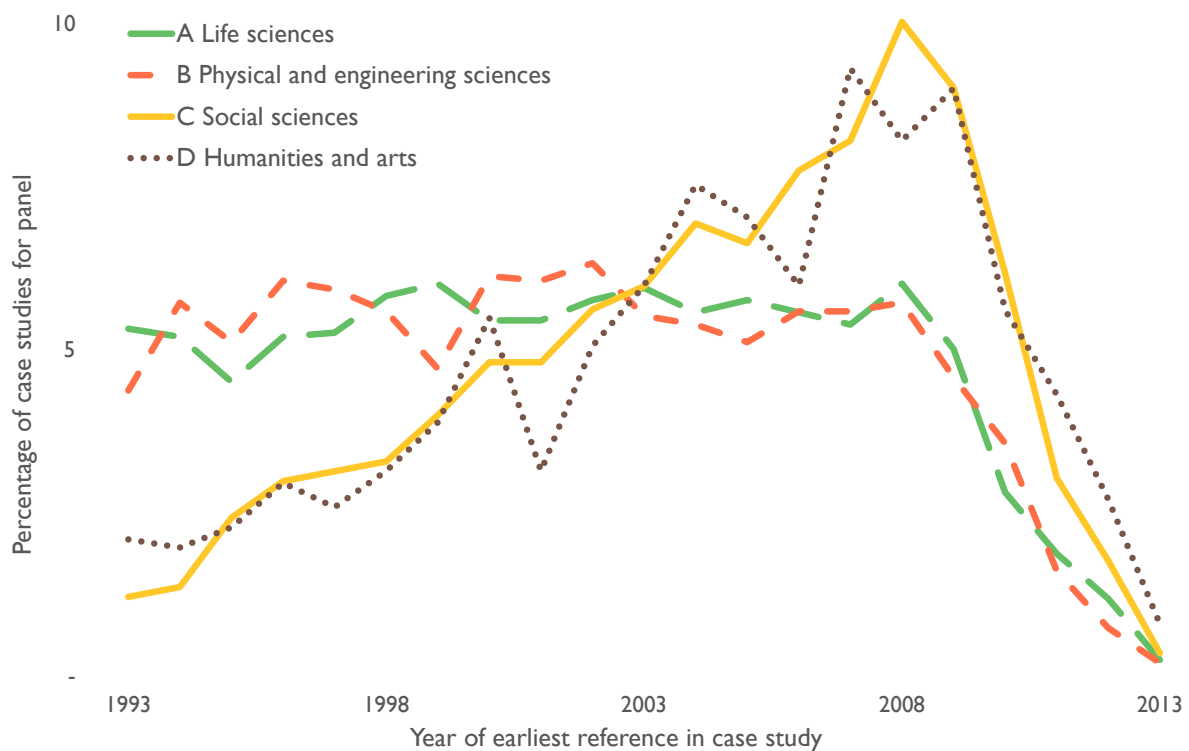


The earliest case study references underpinning impact are evenly distributed by year for science and engineering. An overall time-skew was reported for case study references (King’s College and Digital Science, 2015). Of 36,244 unique references in the case studies, there were 20,779 with publication dates in the same census period as REF2014. Another 11,000 were published in the RAE2008 census period. Fewer than 7,000 (20%) were published before 2003, in the first half of the available window. This skew is similar to that now demonstrated for submitted outputs (Figure 2).

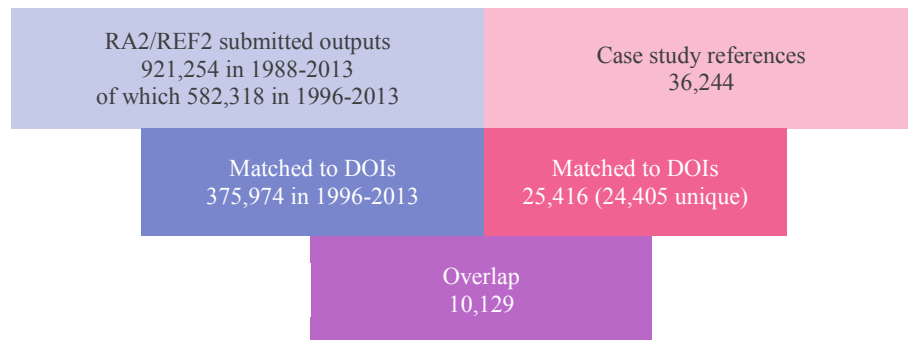
However, impact ‘origin’ is signalled by the earliest reference publication date for each case study. The complete time-spread for all case study references does not describe the specific

origin of intellectual property (IP) underpinning any particular case study. REF2014 required case studies of impact occurring within the 2008-2013 period, but impact emerges as research develops and matures. Disaggregation of the earliest reference dates in each impact case study reveals a cultural divide akin to that for submitted outputs' data. The earliest dates for science and engineering are spread relatively uniformly across 1993-2008 whereas those for social sciences and humanities are skewed to relatively recent years. (Figure 4)

Figure 4. *The spread by publication year of the earliest of six supporting references in each case study grouped by the four main panels of REF2014; case studies were expected to report impact that occurred during the REF2014 census period from 2008, which explains the sharp drop in earliest references with a later publication date.*

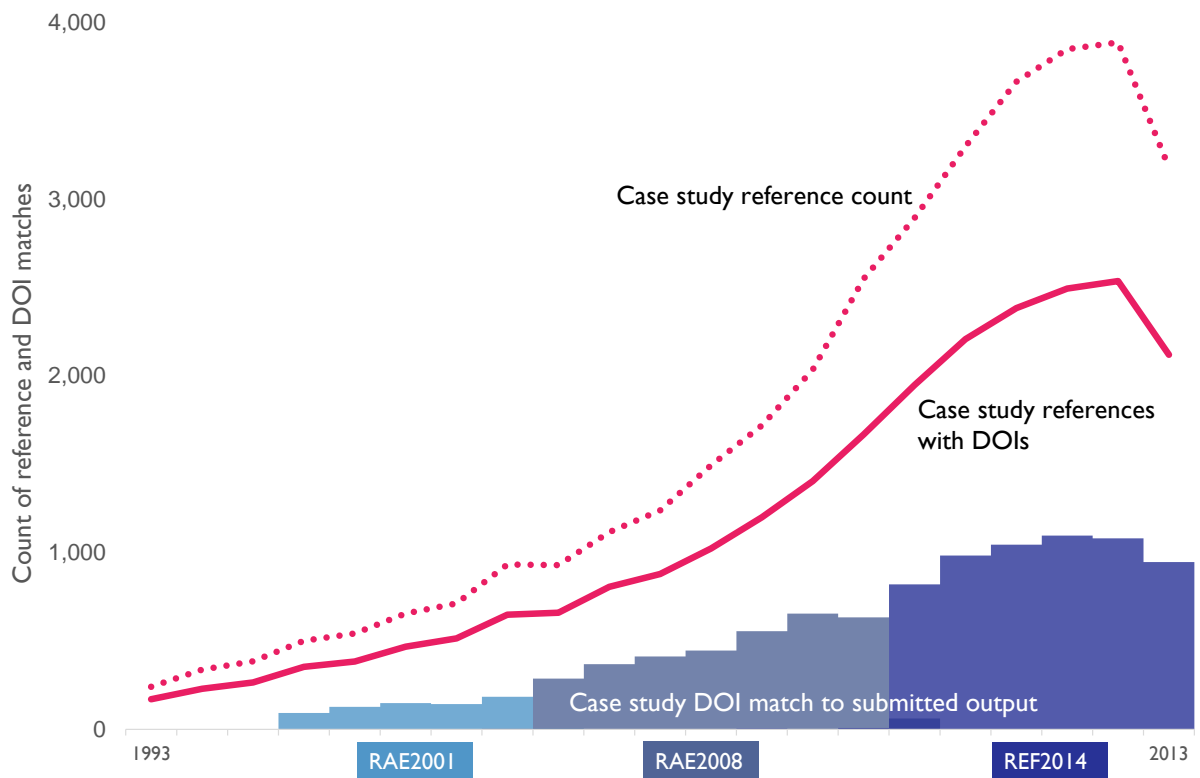


Overlap between submitted outputs and impact references is even across years. The specific overlap between the time-based profiles of the REF2014 case study references and the submitted outputs of that and earlier cycles can be compared for outputs with matched DOIs. Despite early sparsity, and non-journal output types, it was possible to link 25,416 reference records to DOIs. Of these, around 1,000 are duplicates between UOAs or HEIs. This means that there are DOIs for 24,405 (about two-thirds) of the unique case study references.



Analysis of the records matched to DOIs shows that impact case studies drew on outputs previously submitted as evidence of academic excellence at a very consistent rate across 18 years. Overall, 42% of case study references with matched DOIs were also outputs submitted to one of the RAE/REF cycles. For 2001-2013 (the RAE2008 and REF2014 census periods where DOI assignments are comprehensive) the median overlap is 44% and the inter-quartile range is just 43.5 to 45.7. For 1996-2000, sparser DOI matches extend the interquartile range down to 34.3% overlap. (Figure 5)

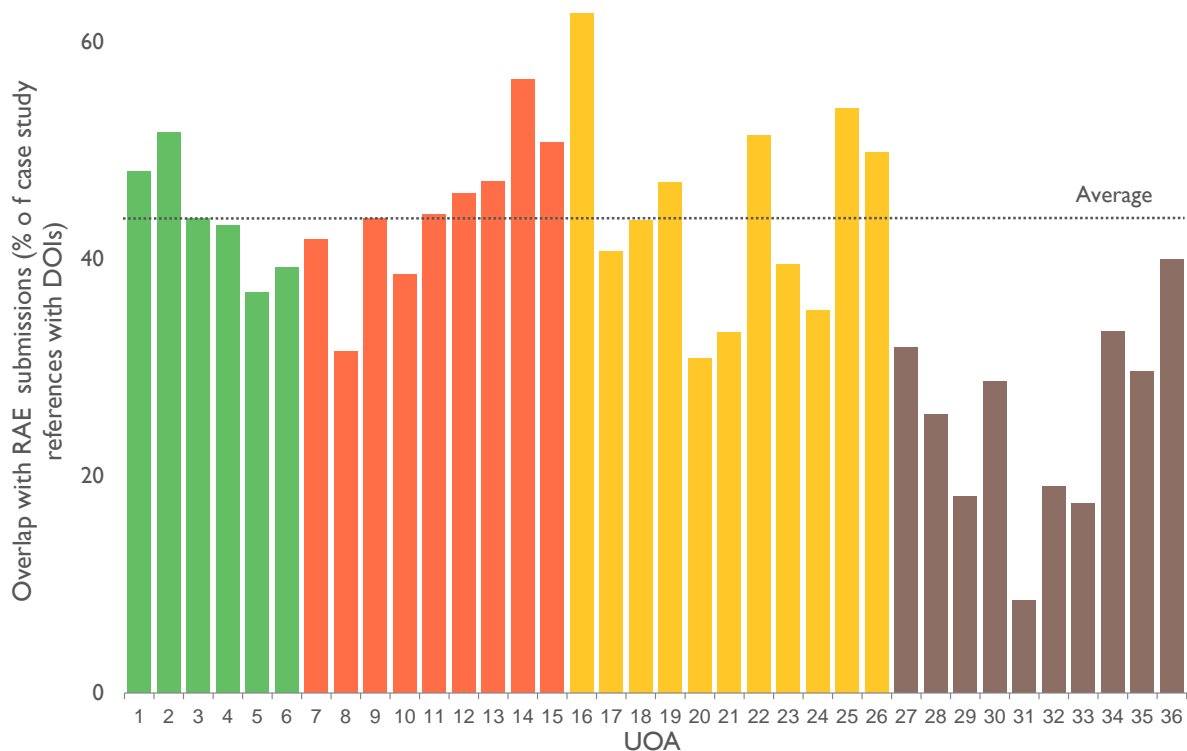
Figure 5. Analysis of publication dates for references included in impact case studies submitted to REF2014. Bars show: the total number of references by publication year; the number of references to which DOIs could be matched; and the number of those references with DOIs that could be paired with outputs submitted in each assessment cycle.



Overlap between submitted outputs and impact references varies by UOA. The average 42% overlap between submitted outputs and references varies by UOA between less than 10% and

over 60% of the case study references. The greatest degree of overlap appears to be in subjects allied to professional practice whilst overlap is less in fundamental research areas in the humanities and core natural sciences. (Figure 6)

Figure 6. *Overlap by UOA (in standard sequence by Main Panel with data mapped to UOAs at 2014) for outputs that could be matched via DOIs. Reference line: average percentage (42%) of case study references also submitted as outputs to RA2/REF2.*



DISCUSSION

This is the first analysis of patterns in the use of published outputs as evidence of research achievement where comprehensive data were available for research published over more than two decades, covering all disciplines and institutions in a national system, with selective submitted outputs to five assessment cycles and for two (complementary but distinct) assessment objectives.

Two characteristics of the selection of submitted outputs progressively changed across RAE cycles and are most readily interpreted with reference to disciplinary differences.

- A distinctive diversity of output types across subject domains in RAE1992, which generally converged on a common preference for journal articles.

Initially scientists favoured journal articles; engineers favoured conference proceedings; and social scientists and humanities researchers submitted more monographs. These differences reinforce the idea of positive selection, presumably driven by cultural prioritisation, and create a reference point of active and discipline-specific choices.

Despite this, the selection of submitted outputs in later assessment cycles shifted almost universally towards journal articles. (Figure 1)

- A marked initial time-skew towards the most recent publication years, which moderated in later cycles.

For RAE1992/96, 30% of the total was selected from just the last full year (1991 or 1995: Figure 2). While there is no obvious reason for academics to select their most recent publications, if each step is progressive then later work absorbs earlier content as well as having novelty appeal. The data indicate strong homogeneity across academic culture: the time-skew is narrowly bounded when disaggregated by discipline and institution. Such homogeneity has been widely supposed but it is rarely evidenced (Bourdieu and Passeron, 1990). In the UK, without discussion or coordination, cognate selection patterns were made across 150 HEIs and 72 UOA subject categories. Despite this, a cultural divide then appears and grows between science/technology and social science/humanities. (Figure 3) What drives these shifts? The RAE datasets may reflect changes in submission choices, in how researchers publish, or a wider shift in research culture to a science model. Journals increased visibility outside the UK, notably in North America, and were part of the growth of international research collaboration (Adams, 2013) and assessment may have accelerated such change. However, a more general influence came from the growing awareness of citation analysis, which links changes in choice of output type and choice of publication date. The first major bibliometric analyses of the UK's comparative international research performance were published after RAE1996 (May, 1997; Adams, 1998). From this, first, journal articles acquired a 'quantitative prioritisation', which may account for changes in submission choices, whilst the moderation and flattening of the time-skew emerges because attention is drawn to more and to less cited articles.

Attention is on socio-economic impact, in case study references, rather than academic impact. The greater diversity of reference output types compared to submitted outputs at REF2014 (Figure 1) thus captures a 'reversion' to the balance seen in RAE1996. There is an overlap between case study references and submitted outputs throughout earlier as well as recent years (Figure 5). The historical output type portfolio may thus have captured a sound balance of impacts, and that is why these publications are recaptured now in impact case studies.

The overlap between references in impact case studies and outputs submitted in prior and current assessment cycles is evidence that work of fundamental quality regularly leads to a wider range of benefits for wealth creation and the quality of life. The overlap accounts for 42% of case study references matched to a DOI, and it is reasonable to infer that a similar overlap would apply to other outputs if they were manually curated. The greater diversity of output types among case study references reflects deliberate selection of outputs as evidence relevant to objective: case study references include more grey literature and patents allied to application and to socio-economic and technological impact.

The overlap also varies by UOA. It is lower in some basic research areas and in the humanities generally, and is higher in areas related to professions. Such divergence may indicate different ways concepts linking research to impact, but further work is required to understand how each discipline constructs this relationship.

There is an interaction between submission profiles and publication dates, for both case study references and submitted outputs. The recency-skew in submitted outputs (Figure 2) was echoed in the overall profile of impact references (King's College and Digital Science, 2015). The skew for impact references in science/engineering disappears, however, when just the earliest 'source' publications are identified (Figure 4). The 20-year period allowed for impact case study references was based on the classic work of Griliches (1986) and Mansfield (1990)

on innovation cycles. The lag between innovation and impact is evidently variable (Figure 4) so the case study reference data only partly support the classic expectation. There have been few authoritative studies on this (but see Wooding et al., 2014) because of the many-to-many relationship between research discoveries and commercially successful outcomes. The time span of original ‘source’ IP in the earliest references supporting case studies covers the full assessment period; it does so at a steady rate for science and engineering; and it seems to interact with impact type and discipline. Drawn-out development might be the case for longitudinal social science research as much as for science, whereas rapid transition to impact may occur for social studies, visual/performing arts and professional health research closely engaged in studying, analysing and reflecting the dynamics of their domains.

REFERENCES

- Adams, J. (1998). Benchmarking International Research. *Nature*, 396, 615-618.
- Adams, J. (2002). Research Assessment. *Science*, 296, 805.
- Adams, J. (2013). The fourth age of research. *Nature*, 497, 557-560.
- Breschi, S., Tarasconi, G., Catalini, C., Novella, L., Guatta, P. and Johnson, H. (2006). Highly Cited Patents, Highly Cited Publications, and Research Networks. CESPRI Report to DG Research, European Commission. http://ec.europa.eu/invest-in-research/pdf/download_en/final_report_hcp.pdf
- Griliches, Z. (1986). Productivity, R&D, and basic research at the firm level in the 1970s. *American Economic Review*, 76, 141-154.
- King’s College London and Digital Science (2015). The nature, scale and beneficiaries of research impact: an initial analysis of REF 2014 impact case studies. Research Report 2015/01, Higher Education Funding Council for England.
- Mansfield, E. (1990). Academic research and industrial innovation. *Research Policy*, 20, 1–12.
- May, R. M. (1997). The Scientific Wealth of Nations. *Science*, 275, 793-796.
- Universities Funding Council (UFC) Circular 5/92 setting out procedures and guidelines for RAE1992 was issued in March 1992.
- Wooding, S., Hanney, S., Pollitt, A., Grant, J. and Buxton, M. (2014). Understanding factors associated with the translation of cardiovascular research. *Implementation Science*, 9, 201-209.



Societal impact metrics for non-patentable research in dentistry ¹

Diana Hicks*, Kim Isett*, Julia Melkers*, Le Song** and Rakshit Trivedi**

*dhicks@gatech.edu; isett@gatech.edu; jmelkers@gatech.edu

School of Public Policy, Georgia Institute of Technology, Atlanta, GA 30332 (USA)

** song@cc.gatech.edu; rtrivedi6@gatech.edu

College of Computing, Georgia Institute of Technology, Atlanta, GA 30332 (USA)

ABSTRACT

Indicators of research impact tend to revolve around patents, licenses and startups. However, much university research is non-patentable and therefore doesn't register in those metrics. That does not mean such research lacks impact, just that it follows different pathways to use in society. Without the visibility of patents, license income and jobs created in startups, society risks ignoring or discounting the societal impact of such research and therefore of undervaluing the research itself. In order to make visible the importance of research advances underpinning broader societal advance, in this project we explore the possibility of developing metrics of research impact for research whose results are relevant to professional practice.

INTRODUCTION

Indicators of research impact tend to revolve around patents, licenses and startups. However, much university research is non-patentable and therefore doesn't register in those metrics. That does not mean such research lacks impact, just that it follows different pathways to use in society. Without the visibility of patents, license income and jobs created in startups, society risks ignoring or discounting the societal impact of such research and therefore of undervaluing the research itself. In order to make visible the importance of research advances underpinning broader societal advance, in this project we explore the possibility of developing metrics of research impact for research whose results are relevant to professional practice. been changed during drafting, e. g. by copying text from a document with different settings.

BACKGROUND

We take as our case study dentistry. Like many professionals – engineers, architects, lawyers etc. - dentists practice solo or in small groups in the community, often without links to centers of research such as hospitals or universities. Professional training aims to develop autonomous judgement, in the diagnosis and treatment of dental patients in this case. Because professional degrees are not research degrees, most professionals do not develop the capacity to assess and absorb research literature, and all are too busy to have the time to do so. Yet university-based dental research continually advances. Realizing an effective connection between research and practice in a timely manner is crucial, not only to justify the public

¹ The authors thank Simone Johnson, Jackie Elliot, Nathanael Pate, Joshua Bowling for research assistance. This research was funded by NICDR grant number U19-DE-22516 within the National Dental Practice Based Research Network (PBRN). Hicks was also supported by NSF Scisip grant 1445121.

money spent on the research enterprise, but to ensure everyone treated by a dentist benefits from current best practice. Professional associations have traditionally played a large role in conveying research advances to practitioners, providing members with practice oriented journals, newsletters, conferences, guidelines and continuing education. Since the 1990s, electronic media have augmented these traditional sources, opening up a multitude of additional channels through which information targeted at practicing dentists is disseminated. These channels take different forms – magazines, blogs, aggregator sites, Wikipedia, Twitter feeds – backed not only by professional associations, but also by for profit publishers, product manufacturers and individual dentists. This presents both challenges for professional associations in reaching dentists as well as new opportunities to study information flow. Social media in health has been examined for its effects on patient understanding, and studies of information in dentistry have simply looked at what dentists read. In contrast, we seek to take advantage of this moment in time to examine the emerging professional knowledge landscape in dentistry to understand how information moves within it, especially how research-based information reaches practicing dentists, and to develop indicators of the influence of research on professional information flows.

The diffusion of research knowledge in dentistry takes place through an increasingly complex and layered set of actors and organizations represented in: academic journals, professional journals, advertising supported magazines, news aggregator websites, sites writing capsule summaries of academic papers, discussion forums, blogs, podcasts, Wikipedia pages and portals, Twitter feeds, LinkedIn sites and Facebook groups. Much of the content outside academic journals re-presents information originated elsewhere. Not only does this create dependencies between channels, but it also implies transformation of the knowledge as it is summarized, interpreted and commented upon. Through this process and in these channels, a filtered subset of research-based knowledge becomes a timely and accessible resource for practicing clinicians.

To understand how research based advances are filtered and move through the range of professional media, we will track the movement of topics through interconnected journals and electronic media and delineate their associated timelines. Assuming that topics that are more widely discussed reach more dentists and are more influential, we will develop metrics of broad, long-lived, clinically relevant discussion in the professional literature and media. Such metrics, when associated with an originating set of journal articles could be used as indicators of research impact.

METHOD

Our study will develop new methods to topic model over time and across a heterogeneous corpus containing examples of every channel directing information at practicing dentists. Computer scientists are currently paying attention to analyzing the spread of concepts, ideas and/or innovations across heterogeneous information resources. In this project, we are particularly interested in topics and information in peer-reviewed journals that motivate information dissemination and discussions in professional space. In new, large online professional communities topics spread among user-generated blogs and forum posts. Sometimes these discussions are inspired by scientific literature. Commercial, ad-supported magazines such as *Inside Dentistry* or *Dentistry Today* also cite scientific journals. These sources also contain discussion of practice management and marketing topics which we will exclude from our analysis. We focus only on clinically relevant material.

We first compile a text corpus including representatives of each type of information sources directed at dentists:

- Research literature as indexed in Web of Science (WoS)
- Professionally oriented journals as indexed in PubMed but not WoS.
- Magazines
- News aggregator sites
- Commercial information sites/blogs
- Blogs
- Discussion forums
- Twitter and other social media
- Wikipedia

Records in the database will contain the time, date, author, title, and abstract (for journal articles) or full text (magazine articles, blog posts, Wikipedia entries etc.) for each item.

Because we examine research impact on professional practice, we exclude from our analysis channels aimed at patients. For example, many dental practices have websites or Facebook pages to provide potential patients with information about the practice or about dental preventive care. In common with the rest of healthcare, dental websites are often aimed at people seeking information relevant to their own health. These are not included here.

Information diffusion across these channels is accompanied by content evolution of the topics, where novel contents, contexts and modalities are introduced by documents in different channels. Topic evolution occurs over time, most document collections are temporal streams and thus the number of topics, the distribution of a topic and topic popularity are time-varying. To examine this, we will acquire the full available archive for each source. In addition, there is a topic hierarchy in such heterogeneous data. For example, while "endodontics" is a topic in itself, it contains many levels of sub-topics like "irrigation" and "intracanal medicaments". After building our collection of documents, and converting each into a set of words carrying content, our first goal will be to discover topics and their hierarchies in the documents across the data sources. Having obtained such a hierarchy, the next goal is to build large scale statistical model to simultaneously track the temporal evolution of any arbitrary topic and reveal the diffusion paths of that topic across heterogeneous resources, with particular focus on topics starting in scientific documents and diffusing through professional outlets.

ANALYSIS

We will select the most prominent 500 topics for in depth analysis and identify where and when each was discussed. This will enable us to ask the following questions:

- How long-lived is discussion of each topic?
- How widespread is discussion of each topic?
- Is there a correlation between longevity and spread of topic discussion?
- What characteristics distinguish widespread and long-lived topics from ephemeral topics?
- What is the share of topics originating outside the corpus, in the dental research literature, in professional journals, or in the dental media?
 - Do any topics originating in magazines or blogs get picked up in the research literature?
- Which topics originating in the research literature are later discussed in the professional literature? Is there a time lag?

- Is there a common trajectory? If so this tells us who watches which outlets which is helpful for those wanting to enhance the diffusion of information to dentists.

With this deeper understanding of the dynamics of topic movement between the research and professional literature, we will develop methods to validate the topics identified and then design metrics to capture success and failure in achieving broad influence in the clinically relevant, professionally-oriented media. Such metrics could be broadly useful across areas of non-patentable research to capture research impact on professional practice in wide swathes of the knowledge economy.



The Evolution of Scientific Trajectories in Rice: Mapping the Relation between Research and Societal Priorities¹

Tommaso Ciarli* and Ismael Rafols**

* t.ciarli@sussex.ac.uk
SPRU, University of Sussex, UK

** i.rafols@ingenio.upv.es
Ingenio (CSIC-UPV), Universitat Politècnica València, Spain;
SPRU, University of Sussex, Brighton, UK;
Observatoire des Sciences et Techniques (HCERES-OST), Paris, France

ABSTRACT

How do specific technologies emerge? How is the scientific research related to technologies prioritised, in governments, firms, universities, and other actors involved in research? Related to these questions is the question of how we investigate trajectories of science and technologies, and how to model their relation. In this paper we aim to answer the initial question for a case study, rice, proposing to use new data, methods, and indicators. We discuss a number of insights on the evolution of the research trajectory on rice since the 1980's, worldwide and for a number of countries which dominate the world rankings in rice production and/or in rice publications. Using a number of socio-economic indicators at the country level, we relate priorities on the research side (publication topics) and on the societal side finding limited relations between supply and demand of research. In the paper we also discuss the combination of scientometric and socio-economic indicators, suggesting that they may be useful for future research on the relation between societal needs and research priorities.

INTRODUCTION

In 1982 Dosi (1982) defined a technology trajectory as “the pattern of normal problem solving activity (i.e. of progress) on the ground of a technological paradigm. [...] Once a path has been selected and established, it shows a momentum of its own [...], which contributes to define the directions towards which the problem solving activity moves. A technological trajectory [...] can be represented by the movement of multi-dimensional trade-offs among the technological variables which the paradigm defines as relevant.” The dynamics of knowledge and technology has been examined from different disciplines – economics, sociology, management and history. Departing from a pure technological explanation, knowledge and technologies evolve in trajectories that are shaped by a variety of techno-economic factors (Freeman 1991), actors (Freeman 1995), socio-economic (Dosi and Nelson 2013; Smith, Stirling and Berkhout 2005) and political (Johnstone and Stirling 2015) factors. The question which remains open is to which extent these different factors contribute in shaping the trajectory of science and technology. The answer to this question is likely to depend on the specific sector, technology, and artefact.

¹ This research has been funded by the UK ESRC (project RES-360-25-0076) and the US National Science Foundation (Award # 1064146). The findings and observations contained in this paper are those of the authors and do not necessarily reflect the views of the NSF.

In this paper we contribute to answering this question by examining the evolution of 30 years of research on rice. We focus on rice because it is an agricultural crop that is a basic good with high technology components, that feeds a huge number of people and that was at the core of the green revolution. Therefore a highly controversial technology, which is likely to be shaped by different factors, in different locations, cultures, and political economies. This paper also aims to contribute to how researchers may answer to the question about which are the main factors shaping knowledge and technology trajectories. We develop scientometric methods to map research priorities in comparison to socioeconomic factors. We map how the research portfolio changes over time and across the main country producers and users of rice crop and publications. We contrast research priorities over time with country achievements (e.g. yield), needs (e.g. consumption and export), and use of inputs (e.g. pesticide use). We discuss the misalignment between research portfolios and characteristics of the main rice producers and users.

Specifically, the paper aims to answer question such as: What are the main research topics in rice research? How have they evolved over time? How has this evolution differed by country? How are research portfolios by country aligned with their socioeconomic needs in terms of yields, nutrition, rice consumption and trade?

In the next section we briefly discuss the background and motivation, we then move to the methods and data, and finally discuss some preliminary and incomplete results.

BACKGROUND AND MOTIVATION

Research prioritisation is one of the aspects shaping knowledge and innovation trajectories. But scientific advance is currently believed to be unevenly distributed across sectors and societal needs (Nelson 2003; Gibbons et al. 1994; Nelson 2011). For example, there is low relative investment in research for diseases affecting poor populations (neglected diseases) and low relative investment in research on healthy lifestyles in comparison to pharmaceutical research (Evans, Shim, and Ioannidis 2014; Røttingen et al. 2013). Technological developments in agriculture privilege specific forms of productivity at the expense of sustainability and tend to neglect local needs (Dalrymple 2006).

Rethorics on societal needs do not always match patterns in science and innovation themselves (Barke and Jenkins-Smith 1993; Plutzer, et al. 1998). This reinforces tendencies for scientific and technological progress to follow paths only partly driven by explicit policy agendas and social needs (David and Sanderson 1997; Miller and Neff 2013; Mokyr 2000).

The chasm between research priorities and societal needs may due to a number of socio economic and political factors. For example, key science policy priorities are typically driven by priorities within the scientific community (Bozeman and Sarewitz 2005), as advocated since Bush (1945) and Polanyi (2000), rather than in response to wider societal needs. Inequality in the distribution of resources and power may be another reason. Different actors have different interests in developing science and innovation and invest in different areas of research portfolios (Wallace and Rafols 2015; Wallace and Rafols 2016)

In order to rebalance research portfolios for science and technology to contribute towards innovation trajectories that better address societal challenges, it is crucial to create better representations and to gather better understandings on the factors and actors that shape science and innovation trajectories.

However, scholarship has faced major methodological difficulties when investigating the relationship between research priorities, innovation trajectories, and societal needs. One of the challenges is the lack of systematic empirical investigation that: i) map current research investments, ii) elicit societal needs and ii) explore their relative alignment. We attempt to address this challenge.

METHODS AND DATA

We analyse data on research output and on rice country statistics. First, the ‘science supply’ analysis is based on publications on rice (105,000 documents) from the CABI database. We use this database for publications, rather than the more common Web of Science and Scopus, because these databases contain a significantly lower number of publications, and are heavily biased against less developed countries (Rafols, Ciarli and Chavarro 2015). This paper is the first large scientometric study exploring the CABI dataset.

Taking a long time window (1983-2012) and using CABI descriptors, we create a global map of rice research which shows the relative size of main topics in rice research (genomics, genetics, pests, weeds, crop yield, socio-economic issues and consumption) and how they are related to each other. We study the evolution of the global portfolios of topics (distribution of publications across topics), and the evolution of country portfolios of topics, for countries with a high proportion of global publications (according to CABI). Figure 1 plots the evolution of country rankings with respect to number of publications from 1983 to 2012 and the total number of publications per year. The top 4 publishers (covering most of the publications in our analysis) are quite stable throughout the period. While below them some interesting dynamics occur, with some countries emerging as relevant publishers, and other fading away with respect to other countries.

In order to map topics and their relations, we use a co-word algorithm to analyse the similarity between keywords. Two descriptors are similar when they appear together in several documents. For example, a paper on genetic sequencing, varietal resistance, and genetic mapping and a paper on genetic sequencing, varietal resistance, genetic mapping and pest control are similar, and they may be on the same topic. Two descriptors are linked when they both co-occur in different documents. We drop descriptors not useful for discrimination because they are too frequent. We use a method based on term frequency-inverse document frequency to identify the most relevant keyword for different subperiods, because an analysis based on the relevance of terms across the whole period of 30 years is likely to bias against terms which are relevant in specific subperiods. Finally, we cluster descriptors defining different topics (communities), and compute the network structure (relation) between descriptors.

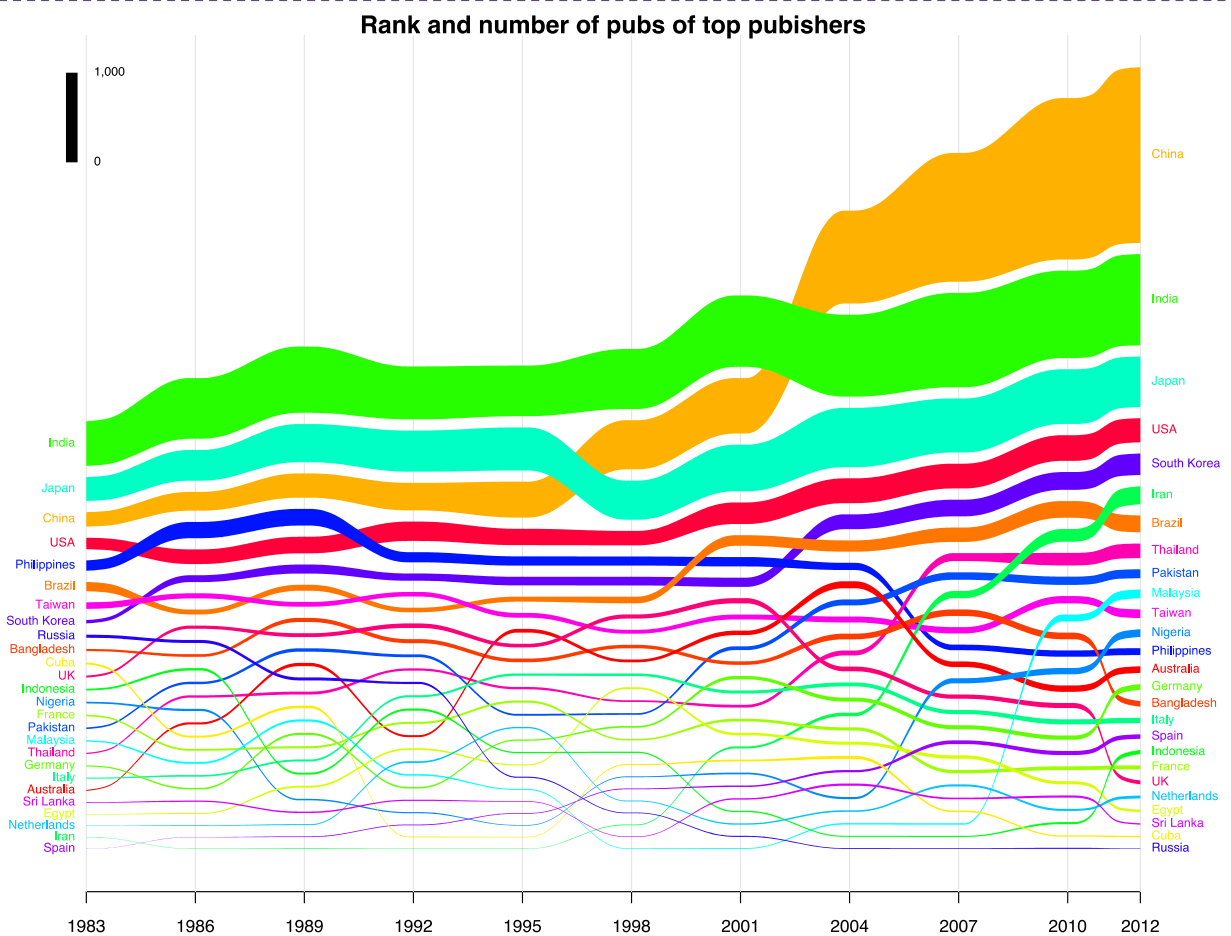


Figure 1 - Changes in the rankings and total number of yearly publications for the top country by number of publications/rice production

Second, we investigate aggregate socioeconomic factors related to the direction of rice research in the countries with the highest rice production and publications on rice. We study how countries' socioeconomic 'demands' or 'needs' change through time, using country data from FAOSTAT on rice cultivation, production, yields, uses (from human nutrition to animal feed) and inputs (pesticides and fertilizer). This allows us to relate countries' characteristics to their research portfolios by investigating the correspondence among relative rankings. We have also run preliminary vector auto-regression models, to study the extent to which the emergence of food, environmental, and commercial issues related to a basic crop influence the direction of research trajectory in a country. Here we present preliminary results. Although some results seem quite robust, we warn the reader that we are still carrying out the analysis, and results may change in future outputs.

PRELIMINARY RESULTS

We first map the global research outputs published between 1983 and 2012 in any country. Figure 2 plots the seven clusters in which the co-wording algorithm groups the publications according to their keywords. Starting from the South (bottom) and moving clockwise these are: 1) Consumption, human nutrition and food technologies such as rice straw, starch, and allergies; 2) Production and socioeconomic issues such as rural development, innovation, and prices; 3) Plant nutrition, such as soil fertility manure and chemical fertilisers; 4) Plant protection against weeds; 5) Plant protection against pests; 6) Rice varieties and classic

genetics such as genetic markers and inheritance; and 7) Transgenics, molecular biology and genomics such as genes or stresses such as heat and salinity.

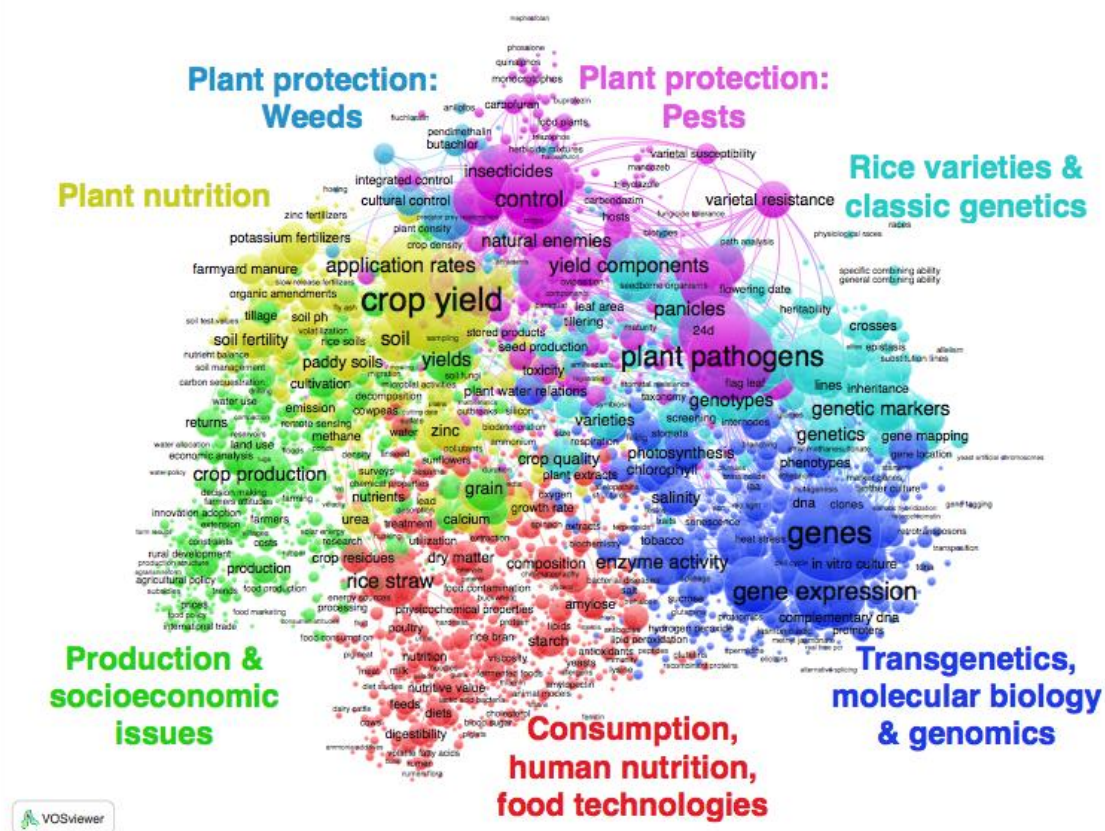
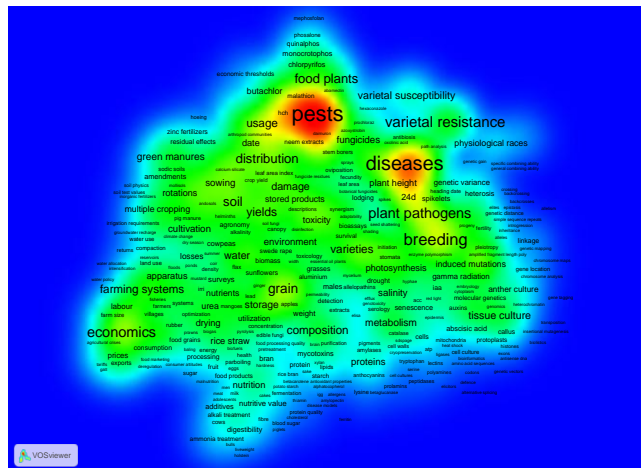


Figure 2 - Global map of rice research (1983-2012). Co-word clustering of publications from the CABI database.

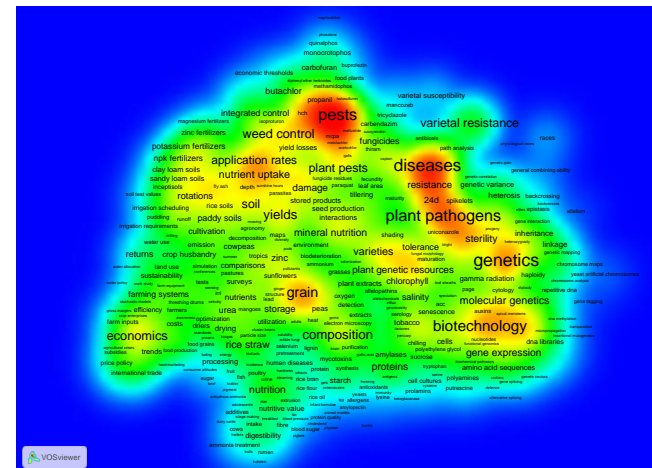
Next, using the base map of descriptors in Figure 2 we overlay (superimpose) maps of subsets of the total corpus of publications to compare which areas of the global map appear with relatively larger frequency. First we investigate how the research trajectory changes through time. Figure 3 plots the relative frequency of descriptors and topics for selected periods (the full distribution is available from the authors): 1983-87 (panel a), 1992-95 (panel b), 2003-04 (panel c), and 2011-12 (panel d). We note that although published research is active in all topics across all periods a number of areas of research emerge during these three decades, while others fade away.

Pests and weeds control is in the 80s the main area of research, as a legacy of the green revolution (panel a). Then yields become acquire relevance, although molecular biology and genomics also start to emerge (panel b). Eventually, molecular biology and genomics become the most active area of research, together with yields, replacing direct research on plant protection at the beginning of the 2000. The other area of research that clearly emerges as relevant is the most recent period is human consumption (south of the map).

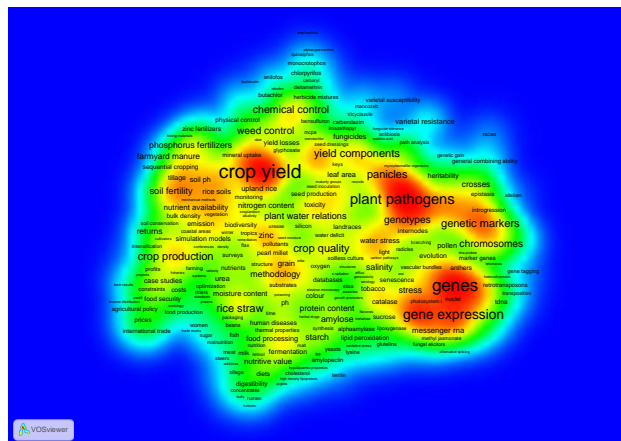
Second, we study how research trajectories differ for different countries among the top publishers and/or producers (see). We focus on the last century only (200-2012). Figure 4 plots frequency maps for four countries for illustrative purposes: India, China. USA and Thailand.



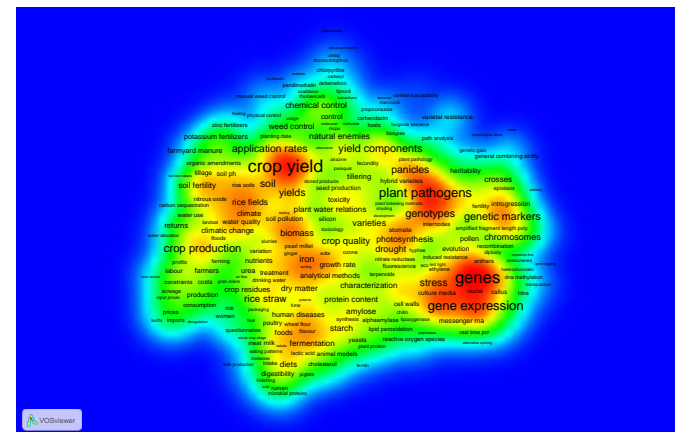
a) 1983-87



b) 1992-95

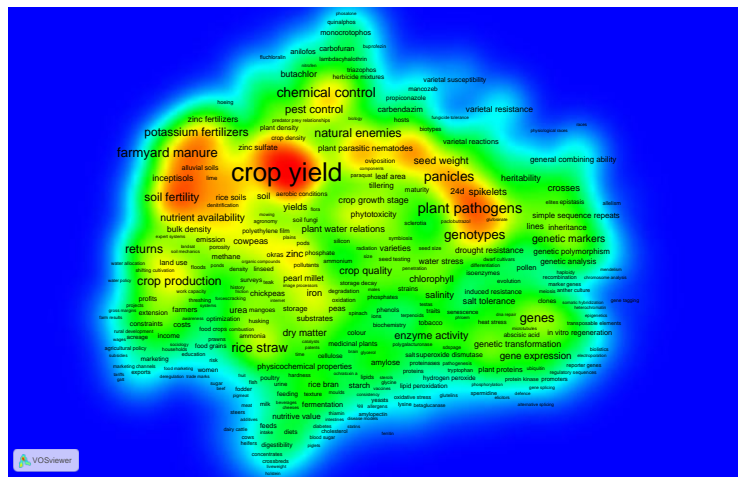


c) 2003-04

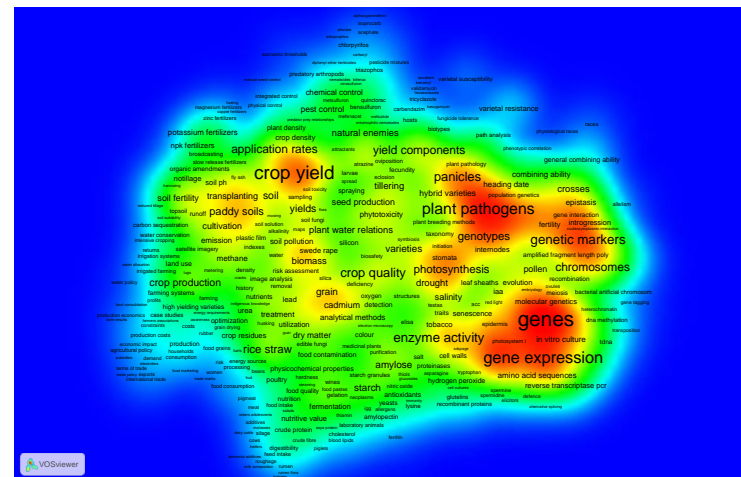


d) 2011-12

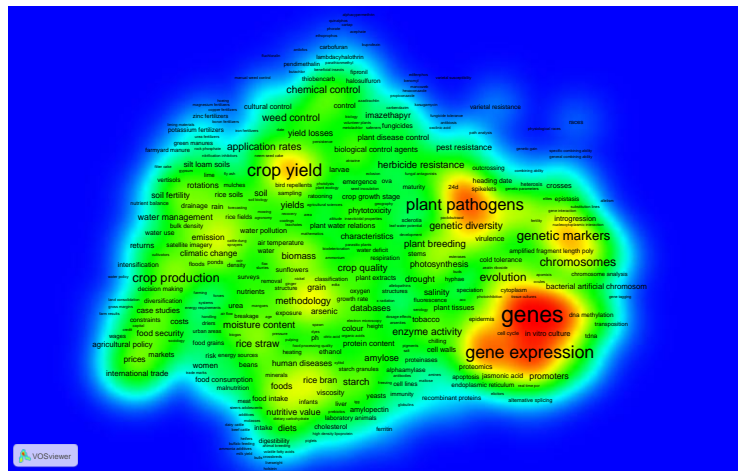
Figure 3 - Frequency of publication for selected subperiods. Heat maps where more red means higher relative frequency and more blue means lower relative frequency. Frequencies are normalised with respect to the maximum and minimum number of appearances of a term in the whole database.



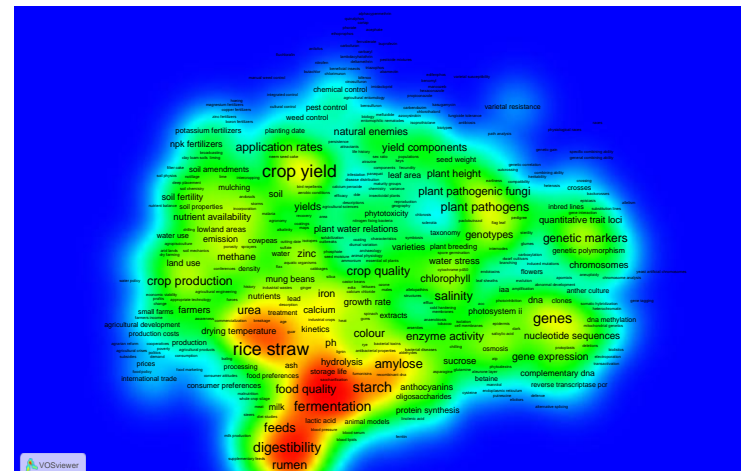
a) India



b) China



c) USA



d) Thailand

Figure 4 - Frequency of publications for selected countries. Heat maps more red means higher relative frequency, and more blue means lower relative frequency. Frequencies are normalised with respect to the maximum number of appearances of a term in a country.



Different countries participate in the global trajectory. Stark differences are found in the contribution between highly industrialised countries such as the US, exporting countries such as Thailand and countries at the heart of the green revolution such as India. The three main global players plotted here (China, India and USA) focus on complementary research. Plant nutrition and protection and traditional breeding in India; different forms of genetic research, and productivity in China; and only transgenic in USA. An exporting country like Thailand mainly does research on human consumption. Indonesia (not shown here), which is the third country by area of rice cultivation, contributes to the trajectory of rice research meagrely in number of publications – ranked 28th (Figure 5).

We then move to discuss preliminary evidence on how the top countries (with respect to rice publications or production) rank on a number of indicators that reflect the country priorities on the demand side, such as consumption, nutrition, trade, use of inputs, and yields. Figure 5 depicts a raking plot. As we have already noticed some top producers, such as Indonesia, rank at the bottom in terms of publications, whereas some top publishers, such as Taiwan or the US, cultivate relatively small areas. It is also interesting to observe the relation between publications and yields. Thailand, the top exporters and among the top ten publishers, experience very low yield. As noted, most of its research is focussed on the human consumption features of rice. India, which focuses most of its research on yields, also performs relatively poorly with respect to yields. However, in those two cases the priorities in publication seem to be aligned with socio economic priorities – exporting for Thailand, and increasing yields for India. In other cases, such as Nigeria, which experiences the lowest yields, and low nutrition, research on socio economic issues rather than yields and consumption may be misplaced.

We are currently carrying out a systematic analysis of these rankings, including the publication topics and the statistical relation between publication and rice output. We therefore postpone further discussion to the conference, when we plan to present more robust results.

LIMITATIONS

This paper has a number of limitations in the empirical strategy adopted that we hope to see discussed and addressed in future work. Mainly, although CABI coverage of publication is possibly the largest on the subject of rice, publications represent a subset of the research. Particularly in agricultural technologies, many research outputs are not accounted for in publications, such as developments on the field, but also a lot of the research done by private companies and public organisations. Secondly, our account of the priorities on the “demand” side is very aggregated (country level), and speaks about revealed needs, not about current needs. A better account and mapping of unpublished knowledge that feed the trajectory and of societal needs is required.

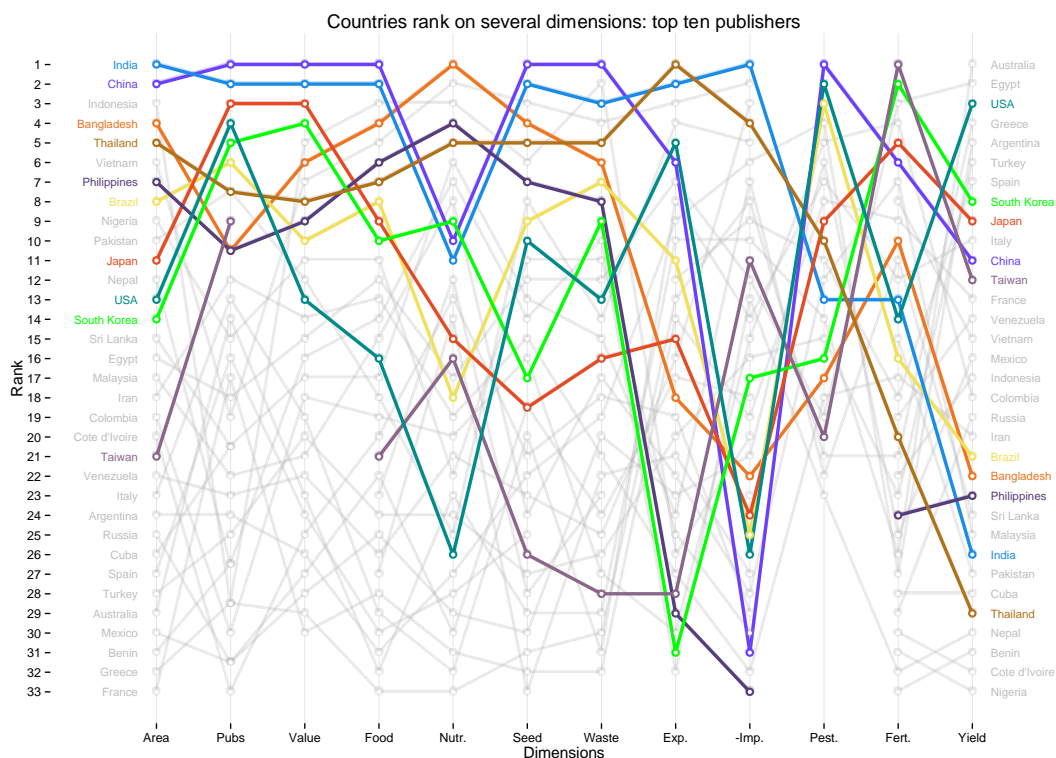


Figure 5 - Countries rank on several dimensions (top ten publishers highlighted).

REFERENCES

- Agarwal, P., and D.B. Searls. 2009. "Can literature analysis identify innovation drivers in drug discovery?" *Nature reviews. Drug discovery* 8(11):865–78.
- Barke, R.P., and H.C. Jenkins-Smith. 1993. "Politics and Scientific Expertise: Scientists, Risk Perception, and Nuclear Waste Policy." *Risk Analysis* 13(4):425–439.
- Bozeman, B., and D. Sarewitz. 2005. "Public values and public failure in US science policy." *Science and Public Policy* 32(2):119–136.
- Bush, V. 1945. "Science: The Endless Frontier." *Transactions of the Kansas Academy of Science* 48(3):231–264.
- Constant, E.W. 1999. "Reliable Knowledge and Unreliable Stuff: On the Practical Role of Rational Beliefs." *Technology and Culture* 40(2):324–357.
- Dalrymple, D. 2006. "Setting the agenda for science and technology in the public sector: the case of international agricultural research." *Science and Public Policy* 33(4):277–290.
- David, P.A., and W.C. Sanderson. 1997. "Making Use of Treacherous Advice: Cognitive Progress, Bayesian Adaptation, and the Tenacity of Unreliable Knowledge." In J. N. Drobak and J. V. C. Nye, eds. *The Frontiers of the New Institutional Economics*. San Diego: Academic Press, pp. 305–366.
- Dosi, G. 1982. "Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change." *Research Policy* 11(3):147–162.
- Dosi, G., and R.R. Nelson. 2013. "The Evolution of Technologies: An Assessment of the State-of-the-Art." *Eurasian Business Review* 3(1):3–46.
- Evans, J. a, J.-M. Shim, and J.P. a Ioannidis. 2014. "Attention to local health burden and the global disparity of health research." *PloS one* 9(4):e90147.
- Freeman, C. 1991. "Innovation, Changes of Techno-Economic Paradigm and Biological

- Analogies in Economics.” *Revue économique* 42(2):211.
- Freeman, C. 1995. “The ‘national system of innovation’ in historical perspective.” *Cambridge Journal of Economics* 19(1):5.
- Gibbons, M., C. Limoges, H. Nowotny, S. Schwartzman, P. Scott, and M. Trow. 1994. *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. SAGE Publications.
- Johnstone, P., and A. Stirling. 2015. “Comparing Nuclear Power Trajectories in Germany And The UK: From ‘Regimes’ to ‘Democracies’ in Sociotechnical Transitions and Discontinuities.”
- Miller, T.R., and M.W. Neff. 2013. “De-Facto Science Policy in the Making: How Scientists Shape Science Policy and Why it Matters (or, Why STS and STP Scholars Should Socialize).” *Minerva* 51(3):295–315.
- Mokyr, J. 2000. “Evolutionary phenomena in technological change.” In J. M. Ziman, ed. *Technological Innovation as an Evolutionary Process*. Cambridge: Cambridge University Press, pp. 52–65.
- Nelson, R.R. 2003. “On the uneven evolution of human know-how.” *Research Policy* 32(6):909–922.
- Nelson, R.R. 2011. “The Moon and the Ghetto revisited.” *Science and Public Policy* 38(9):681–690.
- Plutzer, E., A. Maney, and R. O’Connor. 1998. “Ideology and elites’ perceptions of the safety of new technologies.” *American Journal of Political Science* 42(1):190–209.
- Polanyi, M. 2000. “The Republic of Science: Its Political and Economic Theory.” *Minerva* 38(1):1–21.
- Rafols, I., T. Ciarli, and D. Chavarro. 2015. “Under-Reporting Research Relevant to Local Needs in the Global South. Database Biases in the Representation of Knowledge on Rice.” In *International Conference on Scientometrics and Informetrics*. Istanbul: International Conference on Scientometrics and Informetrics.
- Røttingen, J.-A., S. Regmi, M. Eide, A.J. Young, R.F. Viergever, C. Ardal, J. Guzman, D. Edwards, S. a Matlin, and R.F. Terry. 2013. “Mapping of available health research and development data: what’s there, what’s missing, and what role is there for a global observatory?” *Lancet* 382(9900):1286–307.
- Smith, A., A. Stirling, and F. Berkhout. 2005. “The governance of sustainable socio-technical transitions.” *Research Policy* 34(10):1491–1510.
- Wallace, M.L., and I. Rafols. 2015. “Research Portfolio Analysis in Science Policy: Moving from Financial Returns to Societal Benefits.” *Minerva*.
- Wallace, M.L., and I. Rafols. 2016. “Shaping the Agenda of a Grand Challenge: Institutional Mediation of Priorities in Avian Influenza Research.” SPRU Working Paper Series No. 2016-02, SPRU - Science and Technology Policy Research, University of Sussex.
- Ziman, J.M. 2000. “Commentary: The Republic of Science: Its Political and Economic Theory.” *Minerva* 38(1):21–25.



Research Quality Plus (RQ+) A Holistic Approach to Evaluating Research

Robert K.D. McLean^{*}, Osvaldo N. Feinstein^{**}

^{*}rmclean@idrc.ca

Policy and Evaluation Division, International Development Research Centre, 150 Kent St., Ottawa, K1G 3H9
(Canada)

^{**}ofeinstein@yahoo.com

Complutense University of Madrid, Avda. de Séneca, 2 Ciudad Universitaria, Madrid, 28040 (Spain)

ABSTRACT

The International Development Research Centre (IDRC) strives to fund "excellence" in research - defining excellence as, "methodologically sound, evidence-based, and scientifically robust". But how is the concept of excellence identified and evaluated across the diverse range of research the IDRC supports?

The Research Quality Plus (RQ+) Assessment Framework was developed in order to address these complex issues. It provides a systems-informed approach to defining and evaluating the quality of research, and its positioning for use and impact. It allows tailoring to context, values, mandate and purpose, and can support planning, management and learning processes at any stage in the lifetime of a research project, program or grants portfolio.

This paper presents: 1) the rationale for RQ+ creation, 2) the RQ+ Assessment Framework, and how it can be adapted and used, 3) our early experiences applying it in a series of external evaluations of research quality. The paper is presented by an RQ+ designer and an external user.

INTRODUCTION

Assessing the quality of research is not a new or novel idea. Researchers have long debated the best criteria and means for determining the scientific rigor and significance of empirical studies in the natural and social sciences. What is different is that the concern with research quality has taken on renewed meaning for academic institutions, governments, foundations, nonprofit agencies, and nongovernmental and intergovernmental organizations in light of the intersecting global interests in quality standards, performance measurement, accountability, evidence-based policy and practice, and value for money in research investments.

This set of global concerns has meant a more acute focus on the merits and shortcomings of traditional deliberative (e.g. peer review) and current analytic (e.g. bibliometric) means of evaluating the scientific merit of research (Feller and Stern, 2006; Wilsdon et al., 2015; DFID 2014, Hicks et al. 2015). In turn, it has encouraged a resurrection of interest in the importance of studying knowledge utilization and knowledge exchange processes, drawing specific attention to defining and assessing the impact or benefit of research to society. And, it has resulted in a more profound appreciation of the fact that science and its concern for generating empirically warranted descriptions and explanations of the social and natural worlds can no longer be considered a largely academic enterprise divorced from social concerns.

The International Development Research Centre (IDRC) in Ottawa, Canada, primarily funds and facilitates global South-based research for development. It strives to evaluate its most

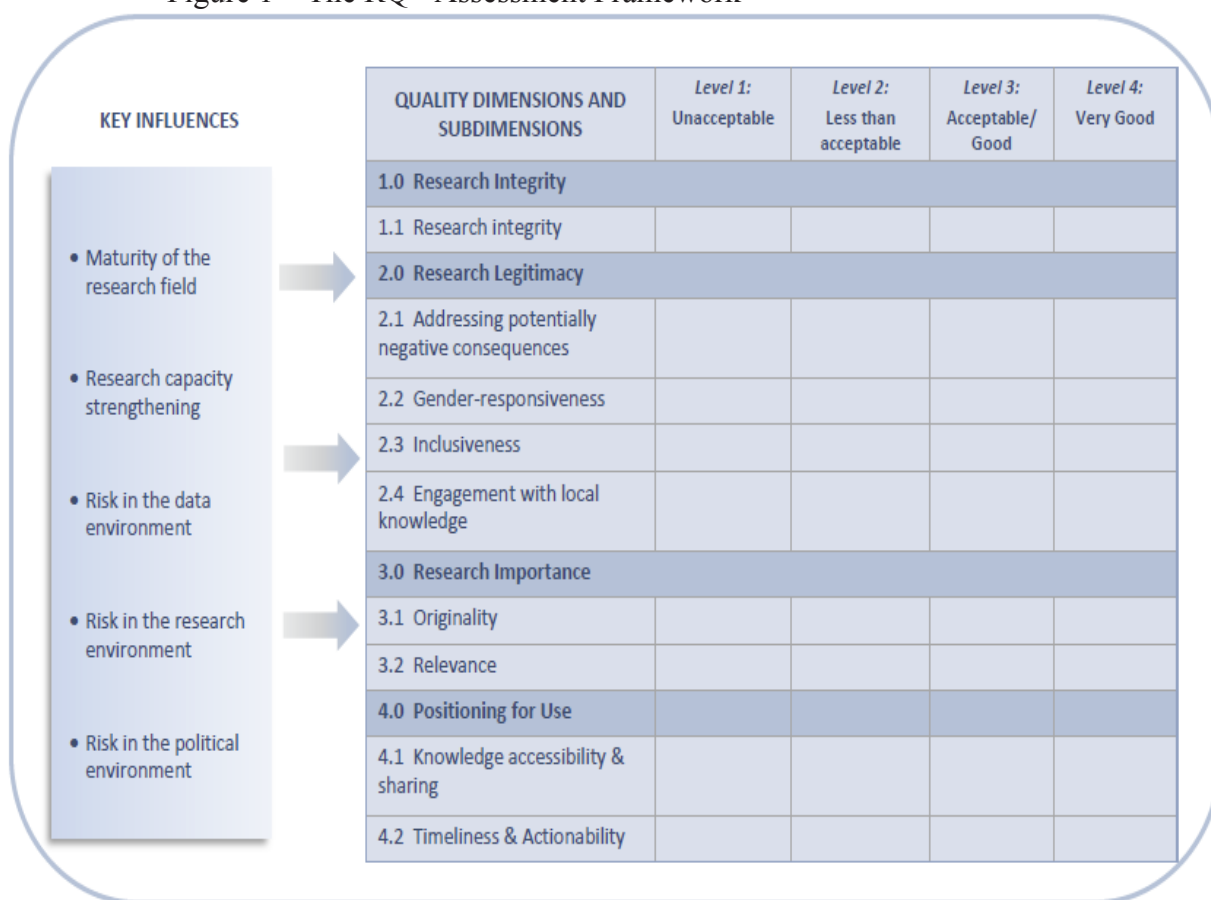
significant investments using numerous evaluative mechanisms, structures, and platforms. Any research assessment framework it employs has to accommodate the heterogeneity of its programs, both in terms of the diversity of *issues* IDRC addresses, such as technology for food security, global health, climate change and inclusive economic growth, and in terms of the types of *activities* it supports, such as research, capacity strengthening, promoting use of research for policy and practice, constituency and network building, and its partnering as a peer, mentor, or broker.

IDRC has launched an effort to develop a new approach to evaluating the quality of the research it funds. The result is presented hereafter as the “RQ+ Assessment Framework”. This work was motivated by IDRC’s desire to advance global research evaluation practice and, more pragmatically, by the need to bring a degree of standardization and transparency to the assessment of research quality – an important emphasis of its evaluative accountability exercise, its External Reviews.

The RQ+ Framework

The RQ+ Assessment Framework consists of the three main components. They include: i) key influences; ii) research quality dimensions and subdimensions; and, iii) evaluative rubrics. The discussion that follows addresses selected aspects of these components in very brief. The following image provides an illustration of the three components and their relationship to one another.

Figure 1 – The RQ+ Assessment Framework



The Key Influences

This component highlights those influences – either within the research endeavor or in the external environment – most likely to affect the quality of the research. Such influences cannot be fully predicted if the assessment is *ex ante*, but this sensitivity to context is one of the most novel aspects of the Framework. The key influences are meant to help evaluators, managers, funders, and others to make meaningful and systematic considerations of the enabling or constraining factors of the research and the risk profile of the project, program, or portfolio, and to incorporate these to the extent possible into their assessments.

For the context in which it operates, IDRC identified five main influences on research quality. **MATURITY OF THE RESEARCH FIELD** - The extent to which well-established theoretical and conceptual frameworks exist and from which well-defined hypotheses have been developed and subjected to testing, as well as a substantial body of conceptual and empirical research in the research field.

RESEARCH CAPACITY STRENGTHENING - The extent to which the research endeavor or project focuses on strengthening research capacities through providing financial and technical support to enhance capacities to identify and analyze development challenges, and to conceive, conduct, manage, and communicate research that can address these challenges.

RISK IN THE RESEARCH ENVIRONMENT - The extent to which the organizational context in which the research team works is supportive of the research, where “supportive” refers, for example, to institutional priorities, incentives, and infrastructure.

RISK IN THE POLITICAL ENVIRONMENT - The extent of external risk related to the range of potential adverse factors that could arise as a result of political and governance challenges, and that could affect the conduct of the research or its positioning for use. These range from electoral uncertainty and policy instability to more fundamental political destabilization, violent conflict, or humanitarian crises.

RISK IN THE DATA ENVIRONMENT - The extent to which instrumentation and measures for data collection and analysis are widely agreed upon and available, and the research environment is data rich or data poor.

The Research Quality Dimensions

There are four principal quality dimensions in the RQ+ Assessment Framework. In our view, these are not discrete. The dimensions are closely interrelated. We have identified and demarcated these dimensions out of a desire for deeper study of the particular issues they represent.

RESEARCH INTEGRITY - Considers the technical quality, appropriateness and rigor of the design and execution of the research as judged in terms of commonly accepted standards for such work and specific methods, and as reflected in research project documents and in selected research outputs. Specified emphases include the research design, methodological rigor, literature review, systematic work, and the relationship between evidence gathered and conclusions reached and/or claims made. Peer reviewed and non-peer reviewed outputs undergo different assessment processes using different criteria.

RESEARCH LEGITIMACY - Considers the extent to which research results have been produced by a process that took account of the concerns and insights of relevant stakeholders,

and was deemed procedurally fair and based on the values, concerns and perspectives of that audience. Legitimacy deals primarily with who participated and who did not; the process for making choices; how information was produced, vetted and disseminated; how well knowledge was localized, and if it respected local traditions and knowledge systems. This dimension also includes a subdimension that asks the assessor to consider the potentially negative consequences and outcomes for populations affected by the research, gender-responsiveness, inclusiveness of vulnerable populations, and engagement with local knowledge.

RESEARCH IMPORTANCE - Considers the importance and value to key intended users of the knowledge and understanding generated by the research, in terms of the perceived relevance of research processes and products to the needs and priorities of potential users, and the contribution of the research to theory and/or practice. Subdimensions include the originality and relevance of the research.

POSITIONING FOR USE - Considers the extent to which the research process has been managed, and research products/ outputs prepared in such a way that the probability of use, influence and impact is enhanced. The uptake of research is inherently a political process. Preparing for it therefore requires attention to user contexts, accessibility of products, and „fit for purpose“ engagement and dissemination strategies. It also requires careful consideration of relationships to establish before and/or during the research process, and the best platforms for making research outputs available to given targeted audiences and users. Positioning for use calls for strategies to integrate potential users into the research process itself wherever this is feasible and desirable. Subdimensions include knowledge accessibility and sharing, actionability, and timeliness.

The first three dimensions – Research Integrity, Legitimacy and Importance – are the core *quality* features typically found in more or less developed forms in most research quality assessment frameworks. The fourth dimension – Positioning for Use – is less typical and is the *plus* (RQ+) feature of the framework. During the Framework development process, IDRC and its research partners determined that it would be reasonable to hold themselves accountable for taking steps to increase the likelihood that the research would be used – in other words, for positioning the research findings for influence and eventual impact.

Figure 2 – The RQ+ Subdimensions



The Evaluative Rubrics

Characterizations for each key influence and performance levels for quality dimensions and subdimensions are based on customizable assessment rubrics that make use of both qualitative and quantitative measures. These narrative behind each rubric is not presented in this paper for sake of brevity but is available in the IDRC full assessment framework guideline (IDRC 2014).

Synthesis of the rubric ratings in different permutations allows for performance to be classified into four levels (from “Unacceptable” to “Very Good”) - as a snapshot of performance, or to follow progress made over time when research capacities are being strengthened and/or projects or programs are on a trajectory towards research excellence.

Figure 3 – A Conceptual Illustration of RQ+



Experience and Lessons from First Application

The Context of First Application

External evaluation is a cornerstone of IDRC's overall evaluation system. External Program Reviews constitute one of the Centre's highest level accountability processes. At the time of this publication, they were structured as summative evaluations of each IDRC program, and conducted once every five years by a team of three independent and external subject area experts, managed by IDRC's Policy and Evaluation Division. The teams of reviewers were comprised of seasoned researchers and evaluators with extensive academic peer review experience.

The remainder of this paper will present IDRC's experience in addressing this issue during the most recent round of External Reviews, conducted between June 2014 and September 2015. The following subsections outline lessons learned in terms of successes, challenges.

Reasons to be Optimistic

Embedded values become explicit - The RQ+ Assessment Framework was created through a reflective process within IDRC. This process aimed to produce assessment criteria which

could encapsulate the organization's values and shared understanding of research quality. This meant the Framework itself guided the external reviewers towards issues that were pertinent, cherished, and specific to IDRC. This Review process, using an instrument that captured the essence of IDRC values, resulted in assessments that primary stakeholders perceived as very useful. The reviewers also benefitted. They felt comfortable making judgments on criteria which were identified as important to the main users, and which IDRC grant recipients were reasonably expected to pursue in their research.

The reviewers noted during their debriefings that the transferability of this aspect of the RQ+ approach held significant potential for application by other research entities. The overall structure and process of implementing the approach could be standard, yet the subdimensions and influencing factors have the potential to change to a lesser or larger extent to reflect the values or learning agenda of another organization. The implementation of the Framework could also have secondary benefits, such as bringing about a common understanding of values within and across organizations, and promoting desired goals in research management and administration.

The systematic nature of the approach strengthened evaluation processes and results - Here we note both process and product optimism. In terms of process, the RQ+ Assessment Framework provided detailed and clear direction to the reviewers. The extensive use of rubrics clarified the basis for the assessments and brought greater consistency to the judgments than typically experienced during conventional peer review processes. The reviewers perceived this as a valuable advantage, despite the significant level of analysis required to use the rubrics across each subdimension and influencing factor. Many reviewers suggested that this "guided process" actually simplified the assessment.

The Framework was also seen as strengthening the result of the External Review. It was given credit for helping to address (albeit not fully) a particular challenge in research evaluation, namely the ability to compare very different types of research. As articulated by one reviewer, the Framework gave "a sense of comfort when comparing the results of a randomized control trial with the results of an advocacy-driven or participatory action research project." Reviewers found this to be a novel and useful contribution. The reasons cited for this advantage included the integration of multiple data sources, the consistent set of dimensions and subdimensions on which all the research was evaluated, and the explicit recognition of the need to consider contextual factors.

The reviewers also expressed their appreciation for the fact that the systematic and detailed process of applying the Framework not only increased their confidence in the quality of their review, but also improved the chance of replication. This is an exciting observation, given the ongoing debates and doubts about the replicability of peer reviews.

Going beyond the research output enables precision, accuracy and richness - The External Review experience demonstrated that the RQ+ Assessment Framework facilitates precise and rich analysis as a result of the comprehensive nature of the approach. This advantage is exemplified in, what reviewers coined, the push to "go beyond the research output." While more complex, it is an improvement over other commonly used peer review systems that are restricted to the object under review. For example, a research paper submitted to a journal is assessed only on the merit of the paper. Likewise, in the case of a research grant application, it is likely that only a research proposal and possibly the track record of the applicant or applicant team will be assessed. In both of these examples, peer reviewers are required to draw predominately, and in most cases only, on their acquired experience and knowledge. It is an unfortunate irony that research evaluation can include so little empirical observation, ie. data collection, validation, and analysis.

To address this, the implementation of the RQ+ Framework requires such peer assessment, but also demands that data be collected from the environment in which the research has occurred. Qualitative interviews with the researchers who conducted the project, research managers, and actual or potential research users (policymakers and practitioners) added richness and resultant accuracy to the assessment at both project and program levels. Other notable methods used by the reviewers to “go beyond the output” included scans of bibliometrics and altmetrics and tailored surveys of “highly influential actors” in the particular field of research.

Challenges and Pitfalls

The assessment is time-consuming, especially when robust triangulation is an imperative - Implementation of the RQ+ approach requires synthesis of qualitative and quantitative data from multiple sources and methods. Extensive consultation with internal and external stakeholders proved to be essential to filling gaps and allowing for sufficient triangulation. The quality and accessibility of sources of project and program data and information are therefore critical factors in the feasibility and value of the approach. Furthermore, where monitoring systems are set up without reference to the Framework components, data collection may require significant time and resources. Reviewers were of the view that the trade-off between comprehensiveness, ease-of-use, and reliable results versus the additional time needed was worthwhile. However, when applying the Framework, careful planning should go into determining a justifiable degree of effort and time.

Quantification after blending quantitative and qualitative data can appear to give simplistic results - At the micro level, the RQ+ approach asks reviewers to assess research projects using both qualitative and quantitative data. Rubrics were considered helpful to bring about more precision in judgment, including by blending the two types of evidence. However, this process became problematic when results were expressed in numerical values (e.g. the rubric ratings). In a sense, without reference to the precise wording of the rubrics, they were perceived as not appropriately capturing the rigor and depth and, hence, the true value and spirit of the assessment. Some reviewers tried to mitigate this perception by using color coding instead of quantitative ratings. The challenge was further compounded by subdimensions that were “not applicable” in certain programs.

At the macro level, data comparison and aggregation presented two challenges: i) understanding the relative values of scores between (sub) dimensions and deciding how these should be weighted and valued, and ii) working with the uncertainties created when following rubric aggregation to the program level, as the Framework guides the synthesis of project assessments into a program-level assessment based on numerical rubric based ratings. The value of a rubric in establishing a program-wide average or composite assessment for influencing factors or subdimensions at an overall program level can be – and was – seen by reviewers and program staff in both positive and negative terms.

Using the RQ+ Assessment Framework for cross-program comparison can have unintended consequences - Directly following from the previous point is the issue of using rubric aggregation as a mechanism for cross-program comparison. For example, if the ultimate user of the assessment is the senior management or board of an organization, such a comparison would be helpful in a strategic, organization-wide assessment of research performance. This can be especially useful if the larger performance discussion is underpinned by the blended qualitative/quantitative nuance that rubrics offer. However, there is a risk that comparison based on the simplistic interpretation of numerical scores might trigger competition among programs and, unwittingly, light the torch of a “program olympics”. This could cause unwarranted and undesirable anxiety among program staff, and even a “race to the bottom” in

performance measures.¹⁷ Such anxiety runs counter to the IDRC's fundamental belief that monitoring and evaluation present useful and constructive methods of self- and external assessment that enable learning, accountability, improvement, and achievement of desirable results. On the other hand, the comparisons could be useful to identify "best practices" and could cause a "race to the peak", which may be used, for example through awards, to encourage emulation and achievement of better results.

The RQ+ Assessment Framework can be used to enrich other assessment frameworks - Although RQ+ was designed for the assessment of the quality of research, it can also be applied in other areas. For this purpose, use can be made of some of the RQ+ research quality dimensions, .e.g., "positioning for use". In this way assessment frameworks such as the United Nations Evaluation Group "Quality Checklist for Evaluation Reports" can be enriched. This is currently being done in the context of an evaluation of FAO's evaluation function. The challenge, however, is to apply the RQ+ to STI and other areas, with appropriate context-derived adaptation. Well adapted it may even be useful to identify key STI indicators.

REFERENCES

- UK Department for International Development. 2014. *What is the Evidence on the Impact of Research for Development?* London, U.K.
- Feller, I. & Stern, P.C., ed. 2006. *A Strategy for Assessing Science: Behavioral and Social Research on Aging*. National Academy of Sciences, Washington, DC.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S. & Rafols I. 2015. "The Leiden Manifesto for Research Metrics." *Nature*, April 23, 2015: 520: 429-431.
- International Development Research Centre. 2014. *Towards Research Excellence for Development: The Research Quality Plus Assessment Instrument*. Ottawa, Canada. Available at idrc.ca/EN/Documents/Research-Quality-Plus-Assessment-Instrument.pdf.
- Wilsdon, J. et al., 2015. *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. Higher Education Funding Council of England, Bristol, UK.



Impact of Research on Development in Cameroon: convergence between supply and research needs in the food sector

Jules René Minkoua* and Ludovic Temple

*minkouarene@yahoo.fr

Department of Public Economics, University of Yaounde II, PO. Box 1365, Yaounde (, Cameroon)

second.author@address.com; m

ludovic.temple@cirad.fr

Cirad, UMR Innovation, 73 rue JF Breton 34398 Montpellier, France

ABSTRACT

The study analyse the convergence of research priorities to development issues especially from the agricultural and food sector in terms of environmental protection and improvement of the business climate, due to weak institutional interconnection devices. Two types of data are used: bibliometric data and a survey of enterprises. Altogether, 1214 and 1708 bibliographic references were generated from WoS and Scopus respectively (from a total of 9146 and 10 557 publications) in Agriculture, Food and Environment for the period 1991-2015 subject to the condition that the author or at least one of the authors is affiliated to an institution based in Cameroon. A total of 317 agro-food companies were identified from the National Institute enterprise's census. The overall message is a relative weak convergence between development priorities identified from perceptions over agricultural and food companies and research works expressed in the scientific literature. This reflects the weak link between research institutions (universities, research centers) and professional milieus that shape entrepreneurship (firms, policy makers) for guidance of scientific production

INTRODUCTION

The renewal of research policies is at the center of political agendas (national, regional ...) to mobilize the economics of knowledge as a lever to reduce technological inequalities and development. In the Least Developed Countries (LDCs) where the ratio of public research investment out of GNP is generally less than 0.4, this renewal is questioned by the inadequacy of conventional indicators of performance evaluation of science and technology (Rafols and al 2015). used in industrial countries. Indeed, these indicators compare national research systems across different variables.

One of the structuring research questions posed by the literature on the conditions of the emergence of national innovation systems in developing countries ((Lundvall, 2014, Touzard and al. 2015, Temple and al. 2016) is to understand how the supply of scientific knowledge meets the demands of the business sector or civil society.

We propose to answer by analyzing the convergence of research priorities to development issues detected from the agricultural and food sector in developing countries, like Cameroon. The assumption is that of a weak convergence between the research outcomes and enterprise

development stakes particularly in terms of environmental protection and improvement of the business climate, due to weak institutional interconnection devices.

METHODOLOGICAL APPROACH

We analyze how research supply meets demand using descriptive statistics and bibliometric from a framework formalized by Sarewitz and Pielke (2007) and McNie, (2007).

The demand for research is approximated by the needs and research priorities indicators empirically developed from indirect responses of agricultural and food business managers.

The supply for research, in turn, is measured by the relevance of scientific publications produced in the fields of Agriculture and Food and Environment (AFE) by academics, researchers, research centers and other institutions in Cameroon, and international. The topics of interest are identified by keywords: "agriculture", "food" and "environment". Since information on patents were not available, these ones have not been retained.

The convergence assessment procedure is done in two steps.

Firstly, we determine the indicators of the relevance of topics addressed in scientific production regarding the improvement of the business climate and the environmental protection. Thematic relevance levels are estimated by VOSviewer from the frequency of occurrence of certain terms in a sufficiently large number of documents. For this, bibliometric analysis examined the titles and abstracts of publications.

The interest on environmental protection and improvement of the business climate is justified in terms of their presumed impact on the development of developing countries. On one hand, the international agenda post 2015 development gives primary attention to the adaptability of production systems over environmental issues (climate change, pollution, natural resource protection, energy transition) that now appear imperative for any country. Moreover, the Strategic Document for Growth and Employment (SDGE) of the Cameroonian Government and various studies by the World Bank set the improvement of the business climate as a prerequisite for development of industries and economic emergence.

In the second step, we examine the correlation between the relevance of the scientific production and business needs. For this, we meet successively the thematic relevance levels associated with improvement of the business climate and environmental protection with indicators of research needs measured in the companies.

The measure of research needs indicators detected from companies is based on the calculation of scores of thematic related to the improvement of business climate on one hand and to the percentages of firm's decisions to invest in the protection of a given environment thematic on the other hand. As for the prioritization by the scores, the arbitrary values are assigned (3, 2.5, 2, 1.5 and 1) in descending order with which each company manager identifies five main themes that constitute barriers to improving the business climate: the sum of the scores for each thematic gives its accorded priority.

The various business climate and environment protection themes are represented graphically. The x-axis and y-axis respectively indicate the levels of relevance and research needs indicator on thematic. It is thus to assess to what extent the relevance of the theme of the "supply for research" is correlated or not to the indicators of "needs or research priorities of

firms." A line of 45 °, dividing the graph into two equal parts is the perfect alignment between research activity and the alleged priorities of research enterprises.

Two types of data are used: bibliometric data and a survey of enterprises. Bibliometric data is gotten from Web of Science (WoS) and Scopus from the Thomson Reuters and Elsevier publishing companies respectively; these two being the main sources of scientific information in the world (Guz & Rushchitsky, 2009). These data serve as proxy for research activity in Cameroon. Data on scientific research output in AFE is collected for the period 1991-2015 subject to the condition that the author or at least one of the authors is affiliated to an institution based in Cameroon (with a valid address as well). In the search of databases, we focus only on *articles* and *reviews* and fine-tune the search using key words such as *agriculture*, *agronomy*, *food*, *nutrition* or *environment*.

Table 1. Scientific publications in AFE in Cameroon from 1991 to 2015

Types of documents	WoS		Scopus	
	Quantity	Percentage (%)	Quantity	Percentage (%)
Articles	1163	95,8	1618	94,7
Reviews	51	4,2	90	5,3
Total	1214	100	1708	100

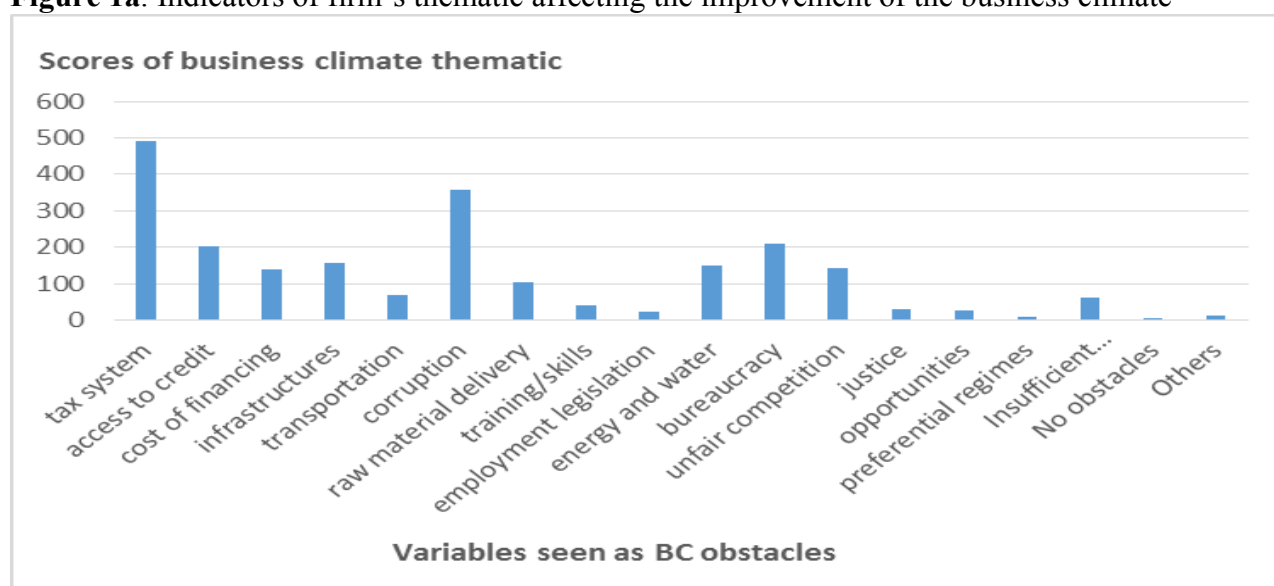
Source: WoS and Scopus data. Construction by authors.

Survey information on enterprises is gotten from the general census of enterprises conducted by the National Institute of Statistics in 2008. A total of 317 agro-food companies were identified

RESULTS AND RECOMMENDATIONS

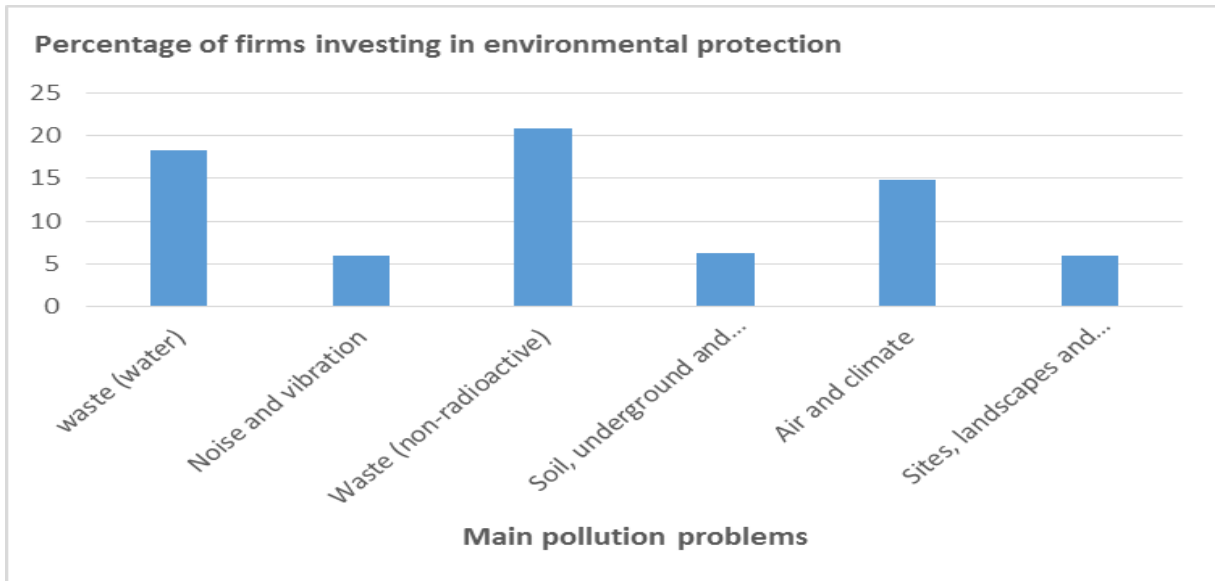
The needs of research perceived from agricultural and food companies on themes concerning the improvement of the business climate and environmental protection are represented by figure 1a and 1b. It shows relatively high scores associated with tax system, corruption, bureaucracy and access to credit and attesting to their high importance for agricultural and food managers. These are the variables that can be seen as obstacles and could relatively either damage or improve the business environment in Cameroon. Other variables have a relatively low importance.

Figure 1a. Indicators of firm's thematic affecting the improvement of the business climate



Source: RGE 2008 (INS). Construction by authors.

Figure 1b. Indicators of firm’s thematic related to environmental protection



Source: RGE 2008 (NIS). Construction by authors.

Concerning environmental protection, only slightly over 20% of companies formally declare investment on topics associated with it. Investment decisions essentially focus on waste management (non-radioactive), recycling of (water) waste and commitments made for the preservation of air and climate.

Regarding the supply for research, publications on AFE (*articles* and *reviews*) contributed 13.3% and 16.2% of all national publication supply respectively in WoS and Scopus. We can see that over the period studied, AFE publications do increase in absolute values, but relatively lower when reporting them to all publications in Cameroon (figure 2a).

Figure 2a. Evolution of publications in AFE

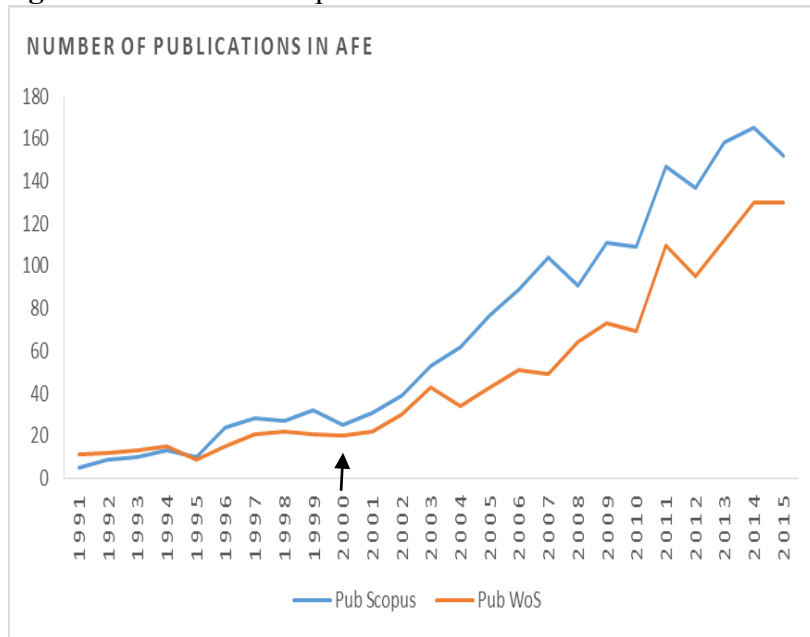
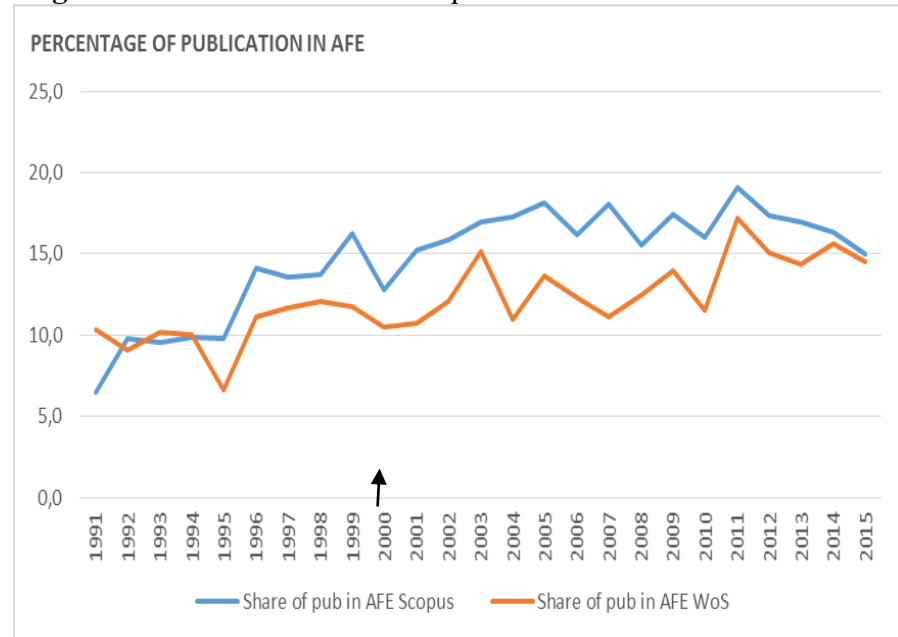


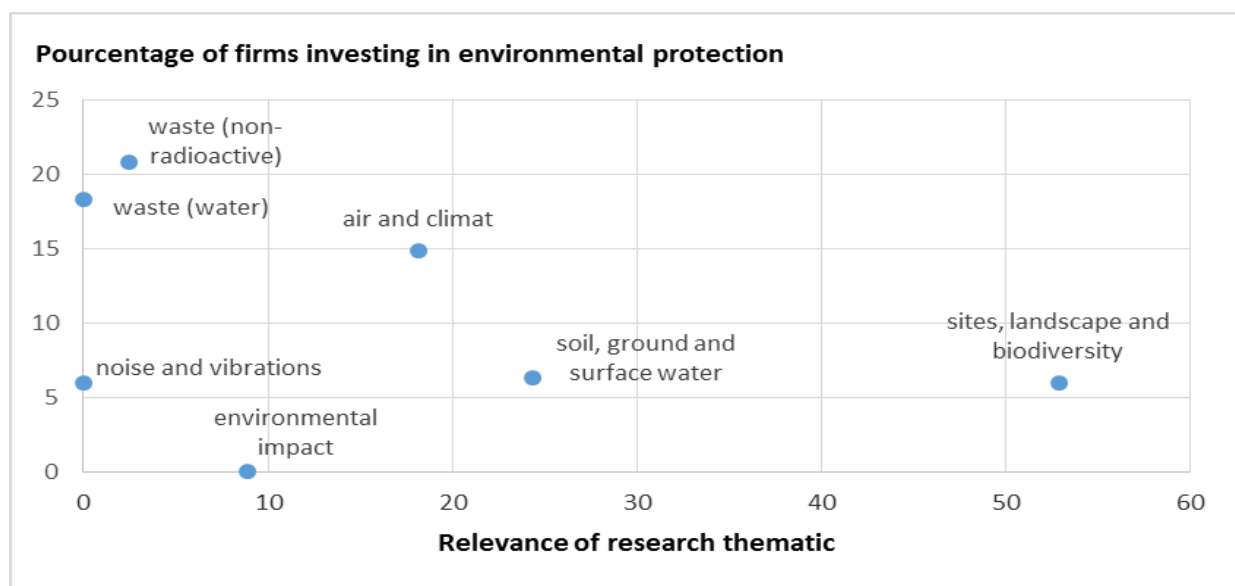
Figure 2b. Evolution of share of publications in AFE



Source: WoS and Scopus data. Construction by authors

It emerges that the balance between the assessment of environmental issues identified by the firms and prevalence of environmental issues addressed by the authors is relatively slim (figure 3). Indeed, if *non-radioactive waste* concerns a relatively large number of companies, this theme and associated expressions have relatively low relevance in scientific studies in Cameroonian production. Contrarily, research topics on *sites, landscapes and biodiversity* have a relatively high pertinent occurrence contrasting however with relatively low perceived importance of these issues at the firm level. It turns out that even some research, especially on the environmental impact, do not arouse the attention of firms, while those that they would be relatively adequate, such as *waste (water)* or *noise and vibrations* are not addressed in the scientific literature. Nevertheless efforts of connection between researchers and companies seem to settle on aspects concerning *air and climate*, and *soil, ground and surface water*.

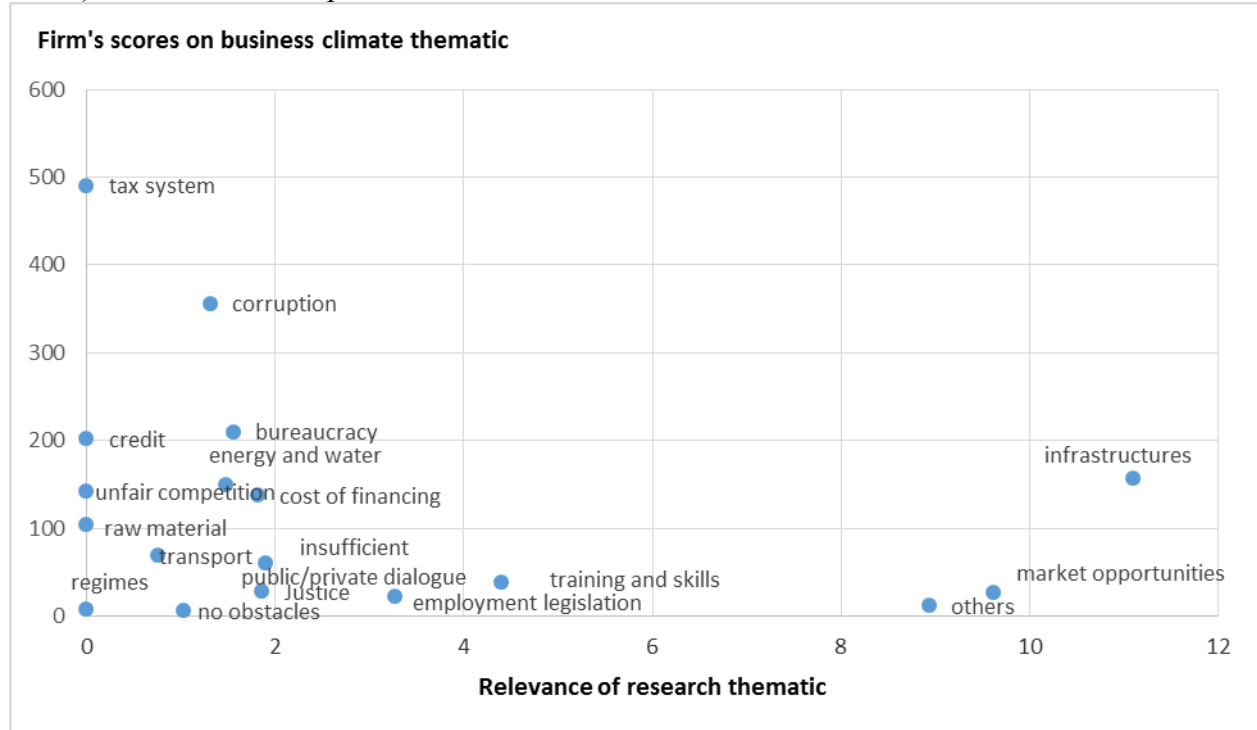
Figure 3. Convergence between supply (thematic relevance) and demand (percentage of investing firms) for research on environmental protection in Cameroon



Source: WoS. Construction by authors.

With respect to variables affecting the improvement of the business climate, the disconnection between the business community and the researchers is striking (figure 4). For example, on important topics such as the tax system, corruption, access to credit, bureaucracy and *water and energy*, scientific work is little or irrelevant. Even the Cameroonian news, especially that relayed by the press and media is quite prolific on these topics. In general, the scientific relevance of indicators are relatively low, as can be noted on important areas of the development process such as *training and skills, infrastructure, and market opportunities*.

Figure 4. Convergence between supply (thematic relevance) and demand (score of firm’s needs) for research to improve the business climate in Cameroon



Source: WoS. Construction by authors.

From the previous results, the overall message is a relative weak convergence between development priorities identified from perceptions over agricultural and food companies and research works expressed in the scientific literature. This reflects the weak link between research institutions (universities, research centers) and professional milieus that shape entrepreneurship (firms, policy makers) for guidance of scientific production (Temple and al 2016). One can notice an extension of the effects of a decade of economic recession and liberalization introduced by the implementation of structural adjustment programs (Cassadela et al 2016), which broke the momentum of the whole economic sectors and institutional arrangements for research and dissemination of innovations. The reconstruction is slow and laborious and one of the challenges is to create a balance between the current structuring of the food sector, dominated by many, almost artisanal production units, with a relative lack of research culture, innovation and the orientation of scientific production (Fofiri et al, 2015). The study, however, raises two questions. The first is on the knowledge of the institutional conditions in which works based on bibliometric analysis can better participate in research policy framing and innovation in developing countries in terms of the results they can produce.

The second highlights the limits of our study based on national survey data that identifies industrial enterprises in the formal sector of the economy. With a total of 77,828 jobs in the industrial sector identified by the 331 surveyed companies and a workforce of around 9,000,000, the industry employs in Cameroon less than 1% of the workforce. Most of the entrepreneurial dynamics that harbor the innovation process are assumed in the informal sector that is to say a complex set of production activities and services related to agriculture

and food activities sometimes in extension of the domestic economy. But the needs and accompanying demand for innovation in the informal sector is very little explained by previous and current studies.

Bibliography

Arvanitis, R., Waast, R., Gaillard, J. (2000). Science in Africa: a bibliometric panorama using database, *Sciencimetrics*, 47, 457-473.

Cassadela V. Liu Z., UZUNIDIS D. (2015), *Développement économique et capacités d'innovation dans la mondialisation*, Collection Innovation Entrepreneuriat et gestion, ISTE.

Fofiri Nzossie, E. J. F., L. Temple et R. Ndjouenkeu (2015). La contribution de la recherche universitaire à la formation d'un système sectoriel d'innovation agro-alimentaire au Cameroun, *Innovations* 47(2), 55-77.

Guz, A. N., & Rushchitsky, J. J. (2009). Scopus: A system for the evaluation of scientific journals. *International Applied Mechanics*, 45(4), 351-362.

Lundvall, BA. and Lemaapages, R. (2014). Growth and structural change in Africa: development strategies for the learning economy. *African Journal of Science, Technology, Innovation and Development*, 6 (5), 455-466. DOI:10.1080/20421338.2014.979660. 455-466.

McNie E. (2007). Reconciling the supply of scientific information with user demands: an analysis of the problem and review of the literature », *Environmental Science and Policy* 10, 17-38.

Rafols R., T. Ciarli, et D. Chavarro (2015). Under-reporting research relevant to local needs in the global south. Database biases in the representation of knowledge on rice, Proceeding Paper presented in the 2015 Globelics Conference, Havana, Cuba.

Sarewitz D. et R. A. Pielke Jr. (2007). The neglected heart of science policy: reconciling supply of and demand for science », *Environmental Science and Policy* 10, 5-16.

Temple L., Machikou Ndzosop N., Fongang Fouepe G.F., Ndoumbè Nkeng M., Mathé S. (2017). Système national de recherche et d'innovation en Afrique : éclairage du Cameroun. *Innovation* (in press)

Temple L., Bienabe E, Barret D, Saint Martin G. (2016). Methods for assessing the impact of research on innovation and development in the agriculture and food sector. *African Journal of Science, Technology, Innovation and Development* (in press).

Touzard J.-M., Temple L., Faure G., Triomphe B. (2015). Innovation systems and knowledge communities in the agriculture and agrifood sector: a literature review. *Journal of Innovation Economics and Management*, 2 (17), 117-142.



Monitoring the Evolution and Benefits of Responsible Research and Innovation (MoRRI) – a preliminary framework for measuring RRI dimensions¹

Niels Mejlgaard*, Susanne Buehrer**, Erich Griessler***, Ralf Lindner**, Nikos Maroulis****, Ingeborg Meijer*****, Viola Peter****, Ismael Rafols*****, Tine Ravn*, Jack Stilgoe*****, Lena Tsipouri*****, Richard Woolley*****, and Angela Wroblewski***

* nm@ps.au.dk; tr@ps.au.dk

Danish Centre for Studies in Research and Research Policy, Aarhus University, Bartholins Allé 7, Aarhus C, 8000 (Denmark)

** Susanne.Buehrer-Topcu@isi.fraunhofer.de; Ralf.Lindner@isi.fraunhofer.de

Fraunhofer Institute for Systems and Innovation Research ISI, Breslauer Strasse 48, Karlsruhe, 76139 (Germany)

*** erich.griessler@ihs.ac.at; wroblews@ihs.ac.at

Institute for Advanced Studies, Josefstädter Straße 39, Vienna, 1080 (Austria)

**** nikos.maroulis@technopolis-group.com; viola.peter@technopolis-group.com

Technopolis Group Belgium, Avenue de Tervuren 188A, Brussels, B-1150 (Belgium)

***** i.meijer@cwts.leidenuniv.nl

Centre for Science and Technology Studies, Leiden University, P.O. Box 905, AX Leiden, 2300 (The Netherlands)

***** i.rafols@ingenio.upv.es; ricwoo@ingenio.upv.es

INGENIO [CSIC-UPV], Ciudad Politécnica de la Innovación, Camino de Vera, Valencia, 46022 (Spain)

***** j.stilgoe@ucl.ac.uk

Department of Science and Technology Studies, UCL, Gower Street, London, WC1E 6BT (United Kingdom)

***** tsipouri@econ.uoa.gr

Department of Economic Sciences, University of Athens, Sofokleous 1, Athens, 105 59 (Greece)

¹ The work leading to this paper has been funded under Service Contract RTD-B6-PP-00964-2013 ‘Monitoring the evolution and benefits of responsible research and innovation’ between the European Union represented by the European Commission and a consortium led by Technopolis Consulting Group Belgium and further consisting of Aarhus University (Denmark), Leiden University (the Netherlands), and Institute for Advanced Studies (Austria) with INGENIO CSIC-UPV (Spain), Fraunhofer ISI (Germany), University College London (United Kingdom), SPRU (United Kingdom), and University of Athens (Greece) as partners.

ABSTRACT

European efforts to 'mainstream' the notion of Responsible Research and Innovation (RRI) is challenged by a lack of organisational and individual incentives, but also by absence of relevant measures of RRI. The European Commission has launched a study with the aim of developing indicators capturing the evolution and benefits of RRI across Europe. This paper presents the preliminary set of indicators identified by the project consortium.

In recent years, the notion of Responsible Research and Innovation (RRI) has emerged in European policy making. From the perspective of the 'Science with and for Society' (SwafS) scheme of the European Commission (EC), the purpose of promoting RRI is "to build effective cooperation between science and society, to recruit new talent for science and to pair scientific excellence with social awareness and responsibility" (EC, 2016).

Building on work by Von Schomberg (2013), the EC defines RRI as a process which allows "all societal actors (researchers, citizens, policy makers, business, third sector organisations etc.) to work together during the whole research and innovation process in order to better align both the process and its outcomes with the values, needs and expectations of European society" (EC, 2016). Conceptually RRI reflects previous strands of activities such as anticipatory governance (Karinen and Guston, 2010) Constructive, Real-Time and other forms of technology assessment (Rip et al., 1995; Guston and Sarewitz, 2002; Grin and Grunwald, 2000), upstream engagement (Wilsdon and Willis, 2004), and value-sensitive design (Friedman, 1996). In operational terms, however, the EC brings dimensions of public engagement, gender equality, science literacy and science education, open access, ethics and governance under the RRI umbrella. These six RRI 'keys' have informed the composition of the work program for SwafS and the featuring of RRI as a cross-cutting issue of Horizon 2020, intended to be embedded across the priorities of the funding programme.

So far, efforts to 'mainstream' RRI across the European research area have been modestly successful (Mejlgaard and Griessler, 2016). Studies indicate significant obstacles, pertaining not least to disincentivizing reward structures at both organisational and individual level (Smallman et al., 2015). While 'pairing' responsibility and scientific excellence is an explicit aim for the RRI agenda, they are in reality often perceived as contradictory demands by the individual scientists or viewed as unequally important concerns by the research performing - and research funding - organisations. While the production of high impact publications is, e.g., considered a core academic activity clearly carrying merit, engaging in public outreach or stakeholder dialogues might easily be considered peripheral activities without straightforward value for the individual scientists.

Moreover, the lack of adequate measures of responsibility in research and innovation further hampers the mainstreaming of RRI. Inability to evaluate, compare, and benchmark 'performance' in terms of RRI at the national as well as disaggregated levels, constitutes a barrier to any revision of reward schemes and dilutes the potential vitality of the organisational or national 'horse race' for high performance in this area. Identification of useful indicators and metrics for RRI might then contribute to bringing issues of responsibility from a peripheral position and closer to the center of activity. Finally, evidence

of ‘benefits’ or impacts of RRI are likely to further contribute to mainstreaming of RRI activities at the level of member states and R&I organisations, and possibly also to increased attention to this cross-cutting issue across the H2020 programmes.

Against this backdrop, the EC has commissioned a study on ‘Monitoring the Evolution and Benefits of Responsible Research and Innovation’ (MoRRI). It contributes conceptual work on RRI, provides extensive exploration of existing metrics capturing RRI, and develops new indicators requiring primary data collection. The first, completed, phase of the indicator development consisted of:

- Review of literature on the six RRI dimensions;
- Assessments of finished and ongoing (inter)national and EC projects to make an inventory on existing and potential indicators for the six dimensions;
- Assessment of data availability within separate dimensions;
- Development of a list of promising indicators, each subsequently evaluated in terms of robustness, richness and RRI relevance;
- Evaluation of the set of promising indicators with regard to EU28 coverage, potential for sustained data collection, and conceptual coverage;
- Final tailoring of a set of 36 indicators covering elements of RRI. The 36 indicators are listed in Table 1 below.

Table 1. Indicators capturing aspects of RRI

Dimension	Indicator	Indicator Title
Gender equality	GE1	Share of RPOs with gender equality plans
	GE2	Share of female researchers by sector (secondary data)
	GE3	Share of RFOs promoting gender content in research
	GE4	Dissimilarity Index (secondary data)
	GE5	Share of RPOs with policies to promote gender in research content
	GE6	Glass Ceiling Index (secondary data)
	GE7	Gender Pay Gap (secondary data)
	GE8	Share of female heads of RPOs
	GE9	Share of gender-balanced recruitment committees at RPOs
	GE10	Number and share of female inventors and authors
Science literacy and education	SLSE1	Science curricula
	SLSE2	RRI related training
	SLSE3	Science communication (secondary data)
	SLSE4	Citizen science
Ethics	E1	Ethics at the level of Universities
	E2	National Ethics Committees (secondary data)
	E3	Research Funding Organizations Ethics Index
Public engagement	PE1	Models of public involvement in S&T decision making (secondary data)
	PE2	Policy-oriented engagement with science (secondary data)
	PE3	Citizen preferences for active participation in S&T decision making (secondary data)
	PE4	Active information search about controversial technology (secondary data)
	PE5	Public engagement performance mechanisms at the level of research institutions
	PE6	Dedicated resources for PE
	PE7	Embedment of PE activities in the funding structure of key public research funding agencies
	PE8	Public engagement elements as evaluative criteria in research proposal evaluations
	PE9	R&I democratization index
	PE10	National infrastructure for involvement of citizens and societal actors in research and innovation
Open access	OA1	Open Access Literature
	OA2	Data publications and citations per country
	OA3	Social media outreach
	OA4	Public perception of Open Access (secondary data)
	OA5	Funder Mandates (secondary data)
	OA6	RPO support structures for researchers as regards incentives and barriers for data sharing
Governance	GOV1	Composite indicator: Governance for responsible research and innovation (secondary data)
	GOV2	Existence of formal governance structures for RRI within RF and RP organisations
	GOV3	Share of research funding and performing organisations promoting RRI

Out of the 36 selected indicators, 13 exploit secondary data while the remaining 23 require primary data collection. Multiple methods will be applied for this purpose. Table 2 provides an overview.

Table 2. Methods for primary data collection

Methods for collecting primary data	Indicators
'Science in Society' actor survey	PE9, PE10
Research performing organisations survey	GE1, GE5, GE8, GE9, SLSE2, SLSE 4, PE5, PE6, OA6, E1, GOV2, GOV3
Research funding organisations survey	GE3, PE7, PE8, E3, GOV2, GOV3
Register-based data	GE10, OA1, OA2, OA3
Qualitative data, interviews/ desk-research	SLSE1

The MoRRI consortium will make use of the STI 2016 Conference as a platform for presenting and discussing the compilation of indicators tapping into the notion of RRI, addressing specifically the peripheral status of RRI compared to other activities in research and innovation; both in the sense of inadequate measures and in the sense of the lower status of this domain of activity. Furthermore, the next steps and challenges of the project, particularly related to the identification of 'benefits' of RRI, will be discussed.

REFERENCES

- European Commission (2016). Science with and for Society. Retrieved March 25, 2016 from <https://ec.europa.eu/programmes/horizon2020/en/h2020-section/science-and-society>.
- Friedman, B. (1996). Value-sensitive design. *Interactions*, Nov-Dec 1996, 17-23.
- Grin, J., & Grunwald, A. eds. (2000). *Vision Assessment: Shaping Technology in 21st Century Society*. Heidelberg: Springer.
- Guston, D. H., & Sarewitz, D. (2002). Real-time technology assessment. *Technology in Society*, 24, 93–109.
- Karinen, R. & Guston, D. H. (2010). Toward anticipatory governance: The experience with nanotechnology. In: Kaiser, M., Kurath, M., Maasen, S., Rehmann-Sutter, C. (eds.) *Governing Future Technologies: Nanotechnology and the Rise of an Assessment Regime*. Dordrecht: Springer.
- Mejlgaard, N. & Griessler, E. (2016). Monitoring RRI in Europe: Approach and Key Observations. In Lindner, R. et al (eds.) *Navigating towards shared responsibility in research and innovation*. Karlsruhe: Fraunhofer Institute for Systems and Innovation Research ISI.
- Rip, A., Misa, T. J. & Schot, J. eds. (1995). *Managing Technology in Society: The Approach of Constructive Technology Assessment*. London: Pinter Publishers.
- Smallman, M., Lomme, K. and Faullimmel, N. (2015). Report on the analysis of needs and constraints of the stakeholder groups in RRI practices in Europe. Retrieved March 25, 2016 from http://www.rri-tools.eu/documents/10184/107098/RRITools_D2.2-AnalysisNeeds+ConstraintsStakeholderGroupsRRI.pdf/83c55909-118c-4cad-b7e4-74d5a770c8a1
- Von Schomberg, Rene (2013). A vision of responsible innovation. In: Owen, R., Heintz, M. & Bessant, J. (eds.) *Responsible Innovation*. London: John Wiley.
- Wilsdon, J. & Willis, R. (2004) *See-through science: why public engagement needs to move upstream*. Project Report. London: Demos.



***“All this grassroots, real life knowledge”*: Comparing perceived with realised concerns of including non-academic evaluators in societal impact assessment¹**

Derrick, GE. *, Samuel, GS. **

*g.derrick@lancaster.ac.uk
Educational Research, Lancaster University, Lancaster , LA1 4YW (United Kingdom)

** g.samuel@lancaster.ac.uk
Educational Research, Lancaster University, Lancaster , LA1 4YW (United Kingdom)

ABSTRACT

New research assessment frameworks that include societal impact criteria, also require the inclusion on non-academic evaluators (users) as part of the assessment panels. Little research has been conducted on how these user evaluators are received by traditionally academic-led panels, and how their presence influences evaluation outcomes. This is especially the case for evaluations including societal impact criteria. This article uses a mixed-methods approach to explore academic-evaluator concerns about the inclusion of user-evaluators in the assessment process. In addition, it explores how their involvement, influenced the outcomes of the evaluation process.

INTRODUCTION

An analysis of evaluative culture includes how reviewers conceive of the relationship between the evaluation process and the power dynamic present in group assessment situations (Dahler-Larsen, 2011; Lamont, 2009; Langfeldt, 2001, 2004). For the peer review of research, the natural academic peer is synonymous with the concept of an “expert” whose presence acts to legitimise the assessment outcomes (Huutoniemi, 2012). Indeed, the status of peer review panellists is explicitly presented as an important guarantor of legitimate assessment outcomes by organisations openly publicising the names and qualifications of reviewers prior to the assessment.

New research assessment frameworks increasingly include new, previously untested criteria such as a consideration of how research is considered excellent outside the academic sphere (Derrick & Samuel, 2016). This challenges the role of academic peer reviewers alone as sufficiently expert for the robust assessment of impact. As a result, these review processes increasingly incorporate user-assessors in traditionally, academically led peer review panels in order to contribute to the panel based assessment of societal impact components of research. However, little research has focused on how these non-academic (user) evaluators are perceived by their academic peers within assessment panels, and non-academic views contribute to the overall panel discussion and assessment outcomes.

¹ This work was supported by the Economic and Social Research Council as part of an ESRC Future Research Leader Grant [grant number ES/K008897/1].

This article uses a mixed methods approach to compare academic evaluator concerns about including user evaluators in the peer review of societal impact, prior to and after the assessment process. In addition, this article explores how anticipated and realised concerns influenced the assessment of the societal impact of research.

METHODS

Research design

The study utilised the in-vitro approach to assessing academic evaluative cultures. The approach is discussed at length in Derrick & Samuel (forthcoming).

Classification of “user” and “academic” evaluators

The codes used in the results below related to the participant’s panel (Main panel=P0; Sub panel 1=P1, etc) and their evaluation responsibilities (Outputs and Impact (OutImp); Impact only (Imp); or Output only (Out)). As per the REF2014 guidelines (HEFCE, 2010), Users were classified as those evaluators with responsibility to evaluation Impact only (Imp), whereas academics were those who had been involved in assessing outputs (Out or OutImp). This classification was confirmed by asking each participant to whether they agreed with their assigned classification prior to each pre-evaluation interview.

Interviews

Interview participants were purposefully sourced from the UK’s Research Excellence Framework (REF2014) Main Panel A evaluation panel. The choice of the REF2014 was because the evaluation panels were charged with the formal, ex post assessment of the societal impact of research.

A total of 215 evaluators were identified and invited to participate in two interviews; one prior to the assessment taking place (Pre-evaluation interviews); and after the assessment was completed (Post-evaluation interviews). The combination of pre- and post-evaluation interviews is key in the utilisation of the in-vitro approach to assessing academic evaluative cultures (Derrick & Samuel, forthcoming). In total, 62 evaluators agreed to participate (53 Academics; 9 Users) in the pre-evaluation interviews, and 55 evaluators returned to participate in the post-evaluation interviews (49 Academics; 6 Users). This resulted in an 88.7% return rate for interviews (92.4% for Academics, and 66.7% for Users).

The interviews were semi-structured where the discussions and cues about the ordering of and structure of the interviews were taken from the interviewee. In pre-evaluation interviews participants were asked the overarching question; “*What value do you see in including evaluators from non-academic/academic* backgrounds in research evaluations?*”; whereas in the post-evaluation interviews; two questions; “*What different definitions of impact did you encounter from other panellists, if any, during the REF process?*” and “*In our previous interviews, participants spoke about how stakeholders/academics* would bring a different perspective to the discussions.*” All overarching questions were followed by a series of pre-designed prompts to further explore emerging themes in more depth and maximising the strength of the qualitative approach adopted in the study.

Interviews were conducted via telephone, skype or face-to-face and were recorded and transcribed for analysis.

Analysis

In the interests of confidentiality, all participant information was coded and entered into Nvivo for analysis. Analysis used an inductive approach to grounded theory. Such approaches use an exploratory style methodology, allowing concepts and ideas to emerge from the data (Charmaz, 2006; Glaser & Strauss, 1968). This method also empirically grounds theorising to data so that abstract conceptualisation can be developed from a close analysis of the data.

Quantitative results

Where possible, the calculation of quantitative data involved taking the code analysed qualitatively (see above), and then manually checking individual scores within Nvivo to determine numbers.

RESULTS

An analysis of participant interviews showed no difference in the proportion of academics and users (54.7 % and 55.6% respectively) who anticipated a “problem” with the inclusion of users and academics in the panel committee assessing impact. Likewise, a similar proportion of academics and users also felt that there would be no problem during the assessment process (34.0% and 33.3% respectively). Interestingly, proportionally more academics felt that the inclusion of users in the assessment panel would be complimentary.

Table 1: Percentage (%) of participants who felt that including users would be a “problem”

	% Academics (n=53)	% Users (n=9)
Yes	54.7 (n=29)	55.6 (n=5)
No	34.0 (n=18)	33.3 (n=3)
Complimentary	17.0 (n=9)	11.1 (n=1)
Don't know	3.8 (n=2)	0.0

A more qualitative analysis of the basis for these anticipated problems in the pre-evaluation interviews showed 4 major concerns. These included concerns that; Users would be swayed by the impact narrative (n=3); Users have insufficient understanding of the research environment (n=12); Users will be more critical of the impact (n=4); and, Users will bring vested interests to the assessment process (n=14).

A comparison of the concerns that participants expressed before the evaluation process, with those concerns realised and reported at the conclusion of the evaluation process are shown in Table 2.

Table 2 shows that all expect one of the anticipated concerns about the involvement of users in the societal impact assessment process, were not repeated in the post-evaluation interviews. The concern that users would be overly critical of the case studies was mentioned by interviewees more in the post-evaluation interviews (9.1%) than in the pre-evaluation interviews (6.4%). It much also be noted that the concern that user's assessment would be overly swayed by the impact narrative was not realised in the interviews after the assessment process.

These concerns are examined in more detail using a qualitative approach in the next sections.

Table 2: Proportion of anticipated problems (pre-evaluation interviews) compared with reported problems (post-evaluation interviews).

	% pre-evaluation interviewees (n=62)	% post-evaluation interviewees (n=55)
Swayed by impact narrative	4.8 (n=3)	0 (n=0)
Insufficient understanding of research environment	19.3 (n=12)	10.9 (n=3)
Act overly critically	6.4 (n=4)	9.1 (n=5)
Bring vested interests to the evaluation	22.6 (n=14)	21.8 (n=12)

Users would be swayed by the impact narrative

In the pre-evaluation interviews, this concern circulated around the issue of whether the users' independence from the academic world would result in being unable or reluctant to provide an objective, evidence based assessment; instead being "swayed" by the impact narrative;

"I would expect a difference on the weight they place on the real quality of the underpinning research that it was offering something genuinely new and has strong, if you like, the evidence of the narrative or persuasive the narrative of causality might be." P2OutImpl

In the post evaluation interviews, no evaluator explicitly linked the potential of being manipulated by the impact narrative to the user evaluators. In fact, contrary to its mention as an anticipated issue (Table 2), in the post-evaluation interviews participants openly acknowledged the role of the narrative as an impact assessment tool; "First of all, everybody agreed that the narrative aspect was an important part" P0OutImp5, particularly when used to counter impact assessment issues such as causality and attribution (Grant, Brutscher, Kirk, Butler, & Wooding, 2010).

Users have insufficient understanding of the research environment

Pre-evaluation, users were perceived to lack an understanding of the academic environment, and therefore would also lack the objectivity valued (perhaps falsely) by academics as essential to peer review. This concern underpinned the feeling of participants that the presence of users would be “disruptive” or that their contributions to panel debates would not be worthy. It was essential that users “...understood the research process well” and “depending on how much experience they’ve got of strictly academic things as opposed to the rest of the world working things, it could be a challenge for some and there could be things that need a lot of explanation for some.” P3OutImp2.

The post evaluation interviews echoed how some evaluators saw this insufficient experience as a drawback for users; “...the other thing that I noticed in the user evaluators is that not having much experience of what the academic environment is really like is a negative.” P4OutImp2, others considered this lack of a research background as an advantage, helping to focus evaluators on assessing the impact rather than the research;

“So we were -- we were very strictly told not to judge the value of the research itself once had reached the two-star threshold...and these folks just weren't interested in that. They were only interested in looking at straight forward at the impact, and they were sometimes useful to divert this away from thing we were sort of interested in.” P4OutImp1

Users will be more critical of the impact

Other concerns of academic evaluators of user evaluators revolved around the different skills users brought to the evaluation; “.... they will come from a specific background, and they will have a specific set of experiences that are associated with that background.” P3OutImp3. In particular, these experiences were viewed as impinging on user objectivity of impact and therefore more critical regarding the value of different impacts; as “...a lot of stakeholders will be, yeah, coming out of the real world, and therefore will probably be more critical about the real impacts” P0OutImp3. Although this perceived increased criticism was seen as an important “part of the dialogue”, and resulting from users being “...much more familiar with that process of taking and using research. P2OutImp4; it was also perceived as potentially framing their approach to impact assessment. This could result in being “over enthusiastic about something or under enthusiastic about something” P4OutImp6, and thereby resulting in an assessment of certain aspects of impact that differed and perceived as more critical than those of the academic evaluators.

After the evaluation process, a larger proportion of participants expressed examples of when users were more critical of the impact (9.1%, n=5), than was anticipated. In these cases, evaluators felt that the users “expected too much from the impact” P3OutImp4; and as a result had to “talk up the user evaluators sometimes” P2OutImp8. However, it must also be mentioned that a larger proportion (10.9%, n=6) of evaluators felt that the users were less critical than the academic evaluators of impact. Academics were said to be “hardnosed”

when it came to impact, and in these situations the users played a positive role and were “able to override the quite often natural tendency of the academics in the subpanels to be extremely judgemental of each other and rather talk each other down.” P0OutImp2. This had an influence of the final marking of the impact;

“...So probably they started off with higher scores but I have to say, by the time we finished, it wasn't their scores that were moved down but the other panellists that were perhaps moved up to meet them.” P1OutImp7

Users will have vested interests

Despite the belief that the differences that users bring to the evaluation panel would enrich the debate between evaluators; “I see why you've got the two bodies, but I think the dynamic in those discussions you will potentially have can be quite interesting” P2OutImp8, there was a degree of caution regarding the motivations of user contributions. These concerns were seen by academics as the result of users not representing academic organisations; “we do have people on there who are not necessarily dedicated researchers. They are representing other bodies” P3OutImp1, and therefore the basis of their discussion contributions would not be made with the interests of the academe in mind;

“I think it's possible that just as we might stakeholders may have a particular agenda because I think sometimes from nominated, let's say, not pressure groups, but groups that could act as pressure groups like [name] is an excellent pressure group and nevertheless that's what it is, it's a pressure group.” P3OutImp2

This was despite many users representing non-academic organisations that have vested interests in academia (e.g. Funding bodies and research charities). However, a high level of caution was reserved for user evaluators from industrial or pharmaceutical organisations which despite providing a perspective which is “...more whole world than academics”, their “understanding of the market forces, for example, which underpin the ability to commercialize certain things. They might well have a slightly different take on things” P5OutImp5.

Another type of vested interest that was noted was the potential favouritism users might show towards impacts that demonstrated a financial or economic impact;

“But the reality is they [users] only see the bottom line and something that could be very, very exciting in science and really maybe have implications down the road, they may say no, this is too far away, we don't see this making money for us now, et cetera.” P0OutImp1

These concerns were further demonstrated during the impact evaluation exercise, through the post-evaluation interviews. However, contrary to what was anticipated these vested interests were synonymous with bringing a broader expertise to the evaluation panel, thereby assisting the assessment;

“...it's a perspective from the users' about what a significant investment is and how serious a commercial company's involvement, I think that was genuinely useful.” P0P1OutImp1

However, some of the concerns regarding the vested interests of the evaluators described in the pre-evaluation interviews were further expressed as realised during the evaluation process, including how users had “*quite fixed views about impact*” P3OutImp8, and favoured economic impacts; “*...the end user is only useful if it’s useful to them or their company or wherever they came from*” P5OutImp5. This, as one evaluator mentioned was “*...a different perspective to impact, but it’s a perspective that doesn’t help the higher education sector*” P6OutImp2. In addition, contrary to the anticipated vested interests of user-evaluators, academic evaluators themselves noted that they were not free of their own professional interest in the assessment outcome; “*...because we’ve got too much of a vested interest in making sure it [societal impact of research] looks good*” P6OutImp2.

Academics changing their mind about users

An analysis of the post evaluation interviews also showed that 10.3% (n=3) of evaluators who had anticipated a problem with users, changed their mind at the evaluation’s conclusion. For these evaluators, the realisation that users were “*...oddly, well, they were pretty enthusiastic about impact*”; and “*...did bring some additional knowledge*” P1OutImp7. In particular, these evaluators’ fears about the differences in opinion about impact between themselves and users were, in their opinion, not realised;

“In reality, the discussions went on. Once the baseline science had been established the discussions that went on were collegial, and the different perspectives would lead into a consensus. So the sharp distinctions, as some anticipated, at least to the panel I was on, didn’t arrive.” P2Imp1

DISCUSSION

The results show that many of the problems academic-evaluators anticipated about the inclusion of user-evaluators in the panel assessment of societal impact were not realised in their full form in the post-evaluation interviews. These concerns demonstrated the doubt that academics had on the capabilities of users to assessment societal impact within peer review panels (Insufficient understanding of the research environment). In particular, the ways in which users would be capable of providing an objective assessment of impact free from the biases of their background, organisation (Bring vested interests to the evaluation) and the subjective construction of the case study (Swayed by impact narrative). Although, academic-evaluators reported that user-evaluators were more critical in their assessment than was anticipated. However, a greater proportion also noted that users acted as positive advocates for the impact, counteracting the “*natural tendency*” of academics to be “*extremely judgemental*” in their assessments.

The major limitation of this study is that, although views were captured and compared pre- and post-evaluation, thereby providing a proxy for the views and opinions discussed during panel debates, they remain solely proxies. Future research is advised to adopt more qualitative, observations of panel discussions in order to extend the analysis presented here.

REFERENCES

- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative research*. Sage Publications Ltd, London.
- Dahler-Larsen, P. (2011). *The Evaluation Society*: Stanford University Press.
- Derrick, G. E., & Samuel, G. (forthcoming). The in-vitro approach: A qualitative methodology for investigating peer review panel debates.
- Derrick, G. E., & Samuel, G. N. (2016). The Evaluation Scale: Exploring Decisions About Societal Impact in Peer Review Panels. *Minerva*, 1-23. doi:10.1007/s11024-016-9290-0
- Glaser, B. S., & Strauss, A. (1968). A.(1967). The discovery of grounded theory. *Strategies for qualitative research*. London: Weidenfeld and Nicolson.
- Grant, J., Brutscher, P.-B., Kirk, S. E., Butler, L., & Wooding, S. (2010). Capturing Research Impacts: A Review of International Practice. Documented Briefing. *RAND Corporation*.
- HEFCE. (2010). REF2014: Panel Criteria and Working Methods. Retrieved from http://www.ref.ac.uk/media/ref/content/pub/panelcriteriaandworkingmethods/01_12.pdf
- Huutoniemi, K. (2012). Communicating and compromising on disciplinary expertise in the peer review of research proposals. *Social Studies of Science*, 42(6), 897-921.
- Lamont, M. (2009). *How Professors think: Inside the curious world of academic judgement*. Cambridge, MA: Harvard University Press.
- Langfeldt, L. (2001). The decision-making constraints and processes of grant peer review, and their effects on the review outcome. *Social Studies of Science*, 31(6), 820-841.
- Langfeldt, L. (2004). Expert panels evaluating research: decision-making and sources of bias. *Research Evaluation*, 13(1), 51-62.

CHAPTER 5

Mission Oriented Research



Using novel computer-assisted linguistic analysis techniques to assess the timeliness and impact of FP7 Health's research – a work in progress report

Vilius Stanciauskas^{*}, Haroldas Brozaitis^{**}, Natalia Manola^{***}, Omiros Metaxas^{****}, and Mike Galsworthy^{*****}

^{*} vilius@ppmi.lt

Public Policy and Management Institute, Vilnius, Lithuania

^{**} Haroldas@ppmi.lt

Public Policy and Management Institute, Vilnius, Lithuania

^{***} natalia@di.uoa.gr

University of Athens, Greece

^{****} omiros@di.uoa.gr

University of Athens, Greece

^{*****} mike_galsworthy@yahoo.co.uk

Independent consultant, Scientists4EU, UK

ABSTRACT

This paper presents the ongoing developments of the ex-post evaluation of the Health theme in FP7 which will be finalised in early 2017. The evaluation was launched by DG Research and Innovation, European Commission. Among other questions the evaluation asked to assess the structuring effect of FP7 Health on the European Research Area and the timeliness of the research performed. To this end the evaluation team has applied two innovative computer-assisted linguistic analysis techniques to address these questions, including dynamic topic modelling and network analysis of co-publications. The topic model built for this evaluation contributed to comprehensive mapping of FP7 Health's research activities and building of a dynamic topic model that has not been attempted in previous evaluations of the Framework Programmes. Our applied network analysis of co-publications proved to be a powerful tool in determining the structuring effect of the FP7 Health to a level of detail which was again not implemented in previous evaluations of EU-funded research programmes.

INTRODUCTION

This paper presents the ongoing developments of the ex-post evaluation of the Health theme in FP7 (PPMI, forthcoming). The general objective of this evaluation was to assess the relevance, effectiveness and efficiency of the actions funded in FP7 Health, measure the efficiency in the use of the resources and the European added-value; assess the impact and achievements of the Health Theme in FP7 against the strategic objectives of FP7, notably the extent to which the projects attained their objectives and achieved intended results and how well the funding has been converted into results and impacts.

The evaluation was structured into 15 evaluation questions, including two questions which required application of computer-assisted analysis techniques:

- Assessment of the collaborative dimension of FP7 Health projects and their structuring effect on the European Research Area
- Assessment of timeliness of FP7 Health's research in comparison with scientific progress of the area

The evaluation started in September 2014 and the projected end date for the evaluation is December 2016. By the time when this paper was drafted (i.e. March 2016) the evaluation team consisting of PPMI, the University of Athens and independent experts had already built a probability based topic model and carried out network analysis of co-publications. The results will be finalised by the end of 2016. This paper presents the key lessons learned of this process.

KEY EVALUATION CHALLENGES

Assessment of the collaborative dimension of FP7 Health projects and their structuring effect on the European Research Area

European Research Area (ERA) refers to the entirety of measures at European level to promote research and development (R&D) and to create a favourable environment for R&D throughout the European Union. ERA is one of the fundamental cornerstones of the European Single Market. During 2007-2013 FP7 was the main instrument at EU level aimed at supporting research and development. Assessment of the structuring effect of FP7 on the ERA has thus been of utmost importance to the European Commission.

Multiple studies and evaluations have attempted to measure this effect through social network analysis (SNA) techniques where participating organisations were used as the unit of analysis. However, this kind of analysis is subject to some weaknesses. In particular, SNA relies on available administrative data which in the case of FP7 projects allow analysis at the organisational level only. 'Traditional' network analysis assumes that all organisations funded under the same project are interconnected with each other. Thanks to a number of large universities which frequently participate in FP7 the European research network appeared to be very well interconnected.

More recently SNA techniques were combined with bibliometric analysis to compute more precise estimates (for example, see Science-Metrix *et al*, 2015). This marks an overall increasing emphasis on using network analysis in combination with bibliometric analysis to analyse co-publication networks. Co-publication networks represent a significant upgrade to the analysis as scientific publications can be used as a good proxy for identifying concrete scientific collaboration and measuring the length of collaboration, lasting effects of networks, creation of scientific societies and structures and other key elements of ERA.

However, while the new methodologies produced better results, the analyses were mainly restricted to the duration of the projects. Little was known about what was happening after the end of EU funding and whether the networks created during FP7 projects were sustained. To the best of our knowledge ex-post evaluation of FP7 Health was the first project where network analysis of co-publications was expanded to cover the time periods both before the start (i.e. were new networks created during FP7) and after the end of the projects (was the cooperation sustained?).

Assessment of timeliness of FP7 Health's research in comparison with scientific progress of the area

To the best of our knowledge no previous study or evaluation attempted to assess the timeliness of FP7 research activities using computer-assisted techniques. The related issues were typically covered through a combination of qualitative research methods, including interviews, case studies and expert judgement. This analysis yielded a number of examples and areas where FP7 research was leading, but the evidence was anecdotal.

Therefore, assuming there will be no major challenges in getting the Commission's approval to use the OpenAIRE data, we are in a good position to use the person-publication-organisation data which will be necessary for the network analysis under this assignment, including its augmentation (e.g., using dynamic topic modelling to define new and emerging scientific fields and the timeliness of the performed research)

APPLIED METHODOLOGY

Assessment of timeliness of FP7 Health's research in comparison with scientific progress of the area

To establish the timeliness of the performed research we will create a probabilistic topic model capturing time evolving research areas and related communities. The primary aim of the model was to produce natural categories of the outputs of FP7 Health projects and benchmark those to the outputs of comparable funding schemes, including the Wellcome Trust in the UK and National Institutes of Health in the US.

An intelligent and scalable probabilistic framework was developed for mining Text Augmented Heterogeneous Information Networks (TA-HINets). The analysis has been based on a powerful unsupervised PTM engine, Inter@ctLDA, designed to discover, annotate and find patterns in large multi-dimensional archives of documents with thematic relationships. Inter@ctLDA extends well established generative probabilistic topic models (PTMs) based on hierarchical Bayesian analysis and Latent Dirichlet Allocation (LDA) that have been successfully used to mine textual content revealing the underlying thematic structures of document collections (see .

A pilot study was first run in early 2015 when sufficient (but not complete) FP7 Health outputs and associated information had been collected. This was so that an expert could assess the output format and quality. There were 150 topics generated. However, these needed an expert to assess the various aspects of key information and derive a clear accurate title for the nature of the topic. This coding was based on key phrases and words alongside a list of the most prominent projects, research funding lines and journals.

Following the piloting phase the full analysis involving FP7 Health, NIH and WT data over 50,000 documents was carried out in late 2015. These outputs were derived from crawling/mining 116 academic repositories and European Commission's databases containing information on FP7 Health's scientific outputs. Using the topic model developed for the evaluation these publications and their text data were grouped into 299 topics and 8 topic categories

Assessment of the collaborative dimension of FP7 Health projects and their structuring effect on the European Research Area

This analysis resulted from an extension of the dynamic model described above which capture topic shifts and time evolving communities. This was a multiple step process that initially involved the identification and content classification of health related publications, and subsequently their clustering based on publication authors. Thus, we performed two types of content analysis: 1) supervised classification using specific already annotated with MESH terms in PubMed publications; and 2) unsupervised mixed membership clustering techniques to discover and annotate documents with thematic information (topics) and identify similar projects or publications and unique author pairs. The latter technique was eventually adopted to produce the results described below.

All available information, including full text publications, related meta-data (e.g., authors, grants, pdb codes, Mesh terms and organizations) and networks (e.g., authorship or citation networks) were analysed using a scalable probabilistic topic modelling engine in order to infer multi modal topics that tie together research areas (described as a distribution over words, mesh terms and pdb codes) with authors. Based on this effort a list of unique author pairs affiliated to particular FP7 Health projects was built. We then took these author pairs built in FP7 Health and mined back their data in the PMC Europe database to the FP6 period (2002-2006) to see if the same authors had cooperated before.

RESULTS ACHIEVED

Assessment of timeliness of FP7 Health's research in comparison with scientific progress of the area

The evaluation project is still ongoing and the final results are expected to be made public in the beginning of 2017. However, as of March 2016 the draft final results have already been produced and it is not expected that these will change substantially before the report is finalised. Should our work be presented in the STI 2016 conference, we would have to ask for the EC's permission to present some of the detailed findings.

Overall, based on the dynamic topic model built for the evaluation we were able to build a detailed and directly comparable map of funded research activities in FP7 Health, NIH and WT. the research activities were grouped into 299 research topics and 8 topic categories. This analysis helped:

- Understand the funding priorities of the three research programmes
- Identify key topics and categories which FP7 Health has been prioritising more/less than NIH or WT
- Time/trend analysis helped identify the fastest growing/shrinking research topics and their respective positions in FP7 Health, NIH and WT
- All this analysis contributed to the assessment of timeliness of FP7 health's research in comparison with the scientific progress of the area. To our knowledge this was the first assessment of this kind in evaluations of the Framework Programmes.

Assessment of the collaborative dimension of FP7 Health projects and their structuring effect on the European Research Area

Based on network analysis of co-publications we established:

- The extent to which scientific collaboration continued in FP7 Health projects after the end of EU funding
- Estimated number of newly built unique author pairs which did not exist prior EU funding (of which continued collaboration after the end of EU funding)
- Estimated share of projects/research activities where EU funding created new and sustained networks
- Estimated share of projects/research activities where EU funding strengthened networks which already existed prior to EU funding
- Estimated share of projects/research activities where EU funding had no sustained structuring effect
- All this analysis contributed to a comprehensive assessment of the structuring effect of FP7 Health on the European Research Area

REFERENCES AND CITATIONS

REFERENCES

Blei, D., Ng, A., & Jordan, M. (2003). Latent Dirichlet allocation. *JMRL* , 993-1022.

Blei, D. (2012). Probabilistic topic models. *Communications of the ACM* , 77-84.

Steyvers, M., & Griffiths, M. (2006). Probabilistic topic models. In T. Landauer, D. McNamara, S. Dennis, & W. Kintsch, *Latent semantic analysis: A Road to Meaning*.

Lawrence Erlbaum. Teh, Y., Jordan, M., Beal, M., & Blei, D. (2006). Hierarchical Dirichlet Process. *J.Am.Stat.Assoc* 101

Public Policy and Management Institute (forthcoming), Ex-post evaluation of the Health theme in FP7, European Commission

Science-Metrix, Fraunhofer ISI and Oxford Research (2015), Study on Network Analysis of the 7th Framework Programme Participation, Final Report



Professional impact of clinical research¹

Gustaf Nelhans*

*gustaf.nelhans@hb.se

Swedish School of Library and Information Science University of Borås. SE 501 90 Borås, Sweden

ABSTRACT

In this study, professional impact is defined as the academic literature that is cited in the literature that is used by professions in order to pursue skilled activities that are specific to their expertise. Specifically, we are focusing on the clinical guidelines that are used in the many health and medical professions that are issued by government bodies at national and international levels to ensure a certain quality level and to make results comparable at the national level.

To date, more than 50.000 references have been identified in about 500 Swedish clinical guidelines issued by the above mentioned governmental bodies in Sweden. Of these, 73 % of the references have been matched to a PubMed id.

The goal of this project is to develop a conceptual and theoretical contribution to the development of indicators for measuring the impact of research outside of the specifically academic literature.

INTRODUCTION

Recently, ‘societal impact’ has become the focus of many national and international evaluation assessments such as the UK REF and the Australian ERA. In Sweden, the Swedish Research Council (SRC) has proposed that a focus on “impact of research beyond academia” (2015, p. 38). The SRC, in a report to the Swedish government, proposes that 15 per cent of the share of national funding should be based on such indicators (*ibid.*). While it is generally considered an important part of research outcomes to evaluate, it is hard to identify quantifiable indicators that cover all aspects of research impact on society. Therefore, for evaluation purposes, impact case studies have been used to evaluate impact qualitatively by assessment panels.

But measuring “impact of research beyond academia” in all its variants – Bornman (2012), for example lists ‘third-stream activities’, ‘societal benefits’ or ‘societal quality’, ‘usefulness’, ‘public values’, ‘knowledge transfer’ as concepts to be used before settling on ‘societal impact’ – is increasingly problematic the more conceptualizations that are introduced. Additionally, when it comes to quantifying societal impact, a large share of attempts have been limited to social media impact, or “altmetrics”, which, arguably does not seem to cover a very broad range of varieties of societal impact that are occurring.

In this paper it is proposed that instead of reaching for all-encompassing indicators of societal impact in all its qualitative different variants, specific aspects of societal impact should be identified. For instance, interest in research by the general public as problematized in the critical studies of *public understanding of science* should not be confused with indicators of impact of research on professional practitioners in different areas of society.

¹ This work was supported by Minso Solutions AB, Sweden

Regarding impact professional practice, two areas of significance have previously been suggested as specifically worthwhile. According to Wilsdon et al (2015), cited references to research in corporate *patents* for innovation and *clinical guidelines* in the health sector could be pursued. Arguably, a third option would be *standards* documents, issued by international and national bodies that provide cited references to research that could be used for citation studies in the same vein in engineering and technical areas. Bornman et al (2015) also proposes that citations to published research in *policy* documents as a relevant source of indicators.

The aim of this ongoing research is to test and validate indicators of impact of clinical research in the literature of the medical and health professions. Such indicators could then be used as indicators that relate clinical research to the activities of the professions that use it.

Traditionally, bibliometric analyses are based on both references and measured citations that are found in any of the established databases that index the scientific literature, predominantly *Web of Science*, and *Scopus*, and to a certain degree, *Google Scholar*. By doing this, much trust is put in the coverage of these databases, and even though coverage is increasing in many areas, it is still true that only certain kinds of research are covered to a high degree (Hicks et al., 2015).

While *academic impact* (here defined as academic research published and cited in source outlets that are indexed in the citation databases) is covered to a high degree (in life sciences and in the natural sciences) or to a certain degree (in the social sciences and the humanities), impact of indexed literature in publications that are not covered by the citation databases is virtually non-existent. In some regards, the opposite is not true, since non-indexed publications could be found by performing a *Cited reference search* in WoS or searching for *Secondary documents* in Scopus.

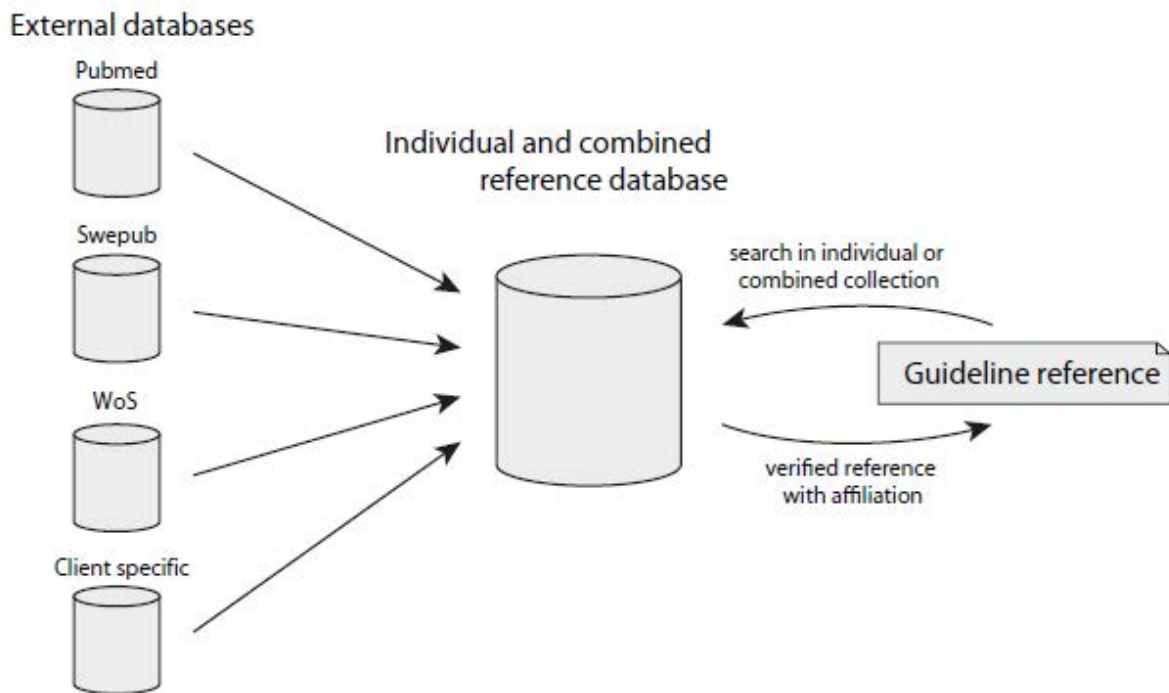
In this study, *professional impact* is defined as the academic literature that is cited in the literature that is used by professions in order to pursue skilled activities that are specific to their expertise. Specifically, we are focusing on the clinical guidelines that are used in the many health and medical professions that are issued by government bodies at national and international levels to ensure a certain quality level and to make results comparable at the national level.

METHOD AND MATERIALS

Examples of such literature are the clinical guidelines that are produced by the *National Board of Health and Welfare* in Sweden, and *SBU – Swedish Agency for Health Technology Assessment and Assessment of Social Services*. These bodies continuously synthesize actual research into guidelines and governing documents for healthcare in practice on a peer review basis. Clinical guidelines has previously been the subject of evaluation of research impact of the medical sciences in the health sector (Andersen, 2013; Lewison, 2004; Lewison and Sullivan, 2008). Here, these results paired with the development of a large database of cited references in Swedish clinical guidelines are used to develop and validate impact indicators at the national level between Swedish regional research bodies within university hospitals.

The process of the research is shown in Figure 1, where the internal database of matched references is validated against external sources such as WoS, assigned with a PubMed id and further validated (if no match is found) against the Swedish national publication database (Swepub) and the possibility of matching against client specific publication databases.

Figure 1. The local reference database based on clinical guideline references verified against Pubmed ID and the local affiliation database.



To date, more than 50.000 references have been identified in about 500 Swedish clinical guidelines issued by the above mentioned governmental bodies in Sweden. Of these, 73 % of the references have been matched to a PubMed id.

Analysis of data

A limited study of cited references from 13 clinical guidelines issued by the Swedish *National Board of Health and Welfare*. Of 6.610 references identified, 5.709 de-duplicated citations were identified by Pubmed ID in Web of Science (Table 1).

Table 1: Clinical guidelines issued by the *Swedish National Board of Health and Welfare*. Name, type of guideline, publication year and number of cited references identified.

Clinical guideline	TYPE	Year	Refs
Cardiac Care	CARDIO	2009	645
Adult Dental Care	DENTA	2011	1301
Diabetes Care	DIABE	2015	523
Methods of Preventing Disease	GENER	2011	316
Antipsychotic Drug Therapy for Schizophrenia or Schizophrenia-type Conditions	MENTH	2014	25
Care in cases of Dementia	MENTH	2010	528
Depression and Anxiety Disorders	MENTH	2010	567
Drug abuse and addiction	MENTH	2015	537
Psychosocial Interventions for Schizophrenia or Schizophrenia-type Conditions	MENTH	2011	106
Treatment of Breast, Prostate and Colorectal Cancers	ONCOL	2014	1.226
Lung Cancer Care and Treatment	ONCOL	2011	277
Palliative Care	PALLIA	2013	78
Musculoskeletal Diseases	REUMA	2012	481
<i>Total</i>			<i>6.610</i>

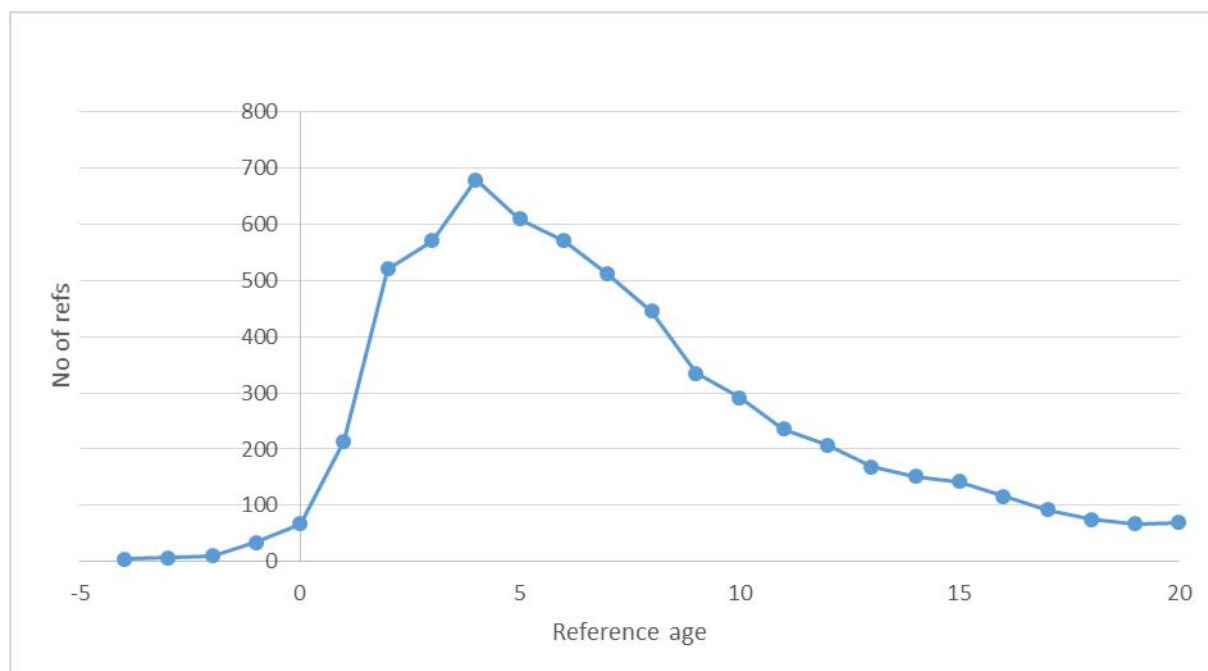
The vast majority of the citations were articles or proceedings papers (86 %), review papers (11.5 %), and other types of literature were in minority.

To a high degree, papers cited in the clinical guidelines are also highly cited in WoS (as indicated by the average number of citations for the ten most cited journals among the guidelines in Table 2). It is noteworthy that this is not always the case. For some journals indicated by italics in the table, the mean number of citations for papers in individual journals were not that high, indicating that to some extent, literature cited in clinical guidelines does not fully overlap with highly cited publications in the academic literature. This is interesting in relation to the results of Thelwall and Maflahi (2016), where it is indicated that since papers cited in NICE clinical guidelines are more cited in general than other papers in the same journal issues, citations in clinical guidelines might not be that different in quality from citations in academic literature. This will be tested further in future research.

Table 2. The ten journals with most cited papers in the clinical guidelines along with total number of WoS citations and mean citations per paper.

#	Journal	Recs	Citations	Avg Cit/rec
1	JOURNAL OF CLINICAL ONCOLOGY	187	54.591	291,9
2	NEW ENGLAND JOURNAL OF MEDICINE	152	16.7521	1102,1
3	STROKE	144	24.435	169,7
4	<i>JOURNAL OF CLINICAL PERIODONTOLOGY</i>	120	7.368	61,4
5	<i>COCHRANE DATABASE OF SYSTEMATIC REVIEWS</i>	112	6.077	54,3
6	LANCET	109	68.433	627,8
7	JAMA-JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION	101	58.214	576,4
8	<i>JOURNAL OF PERIODONTOLOGY</i>	81	3.373	41,6
9	BRITISH MEDICAL JOURNAL	75	18.110	241,5
10	AMERICAN JOURNAL OF PSYCHIATRY	73	11.070	151,6

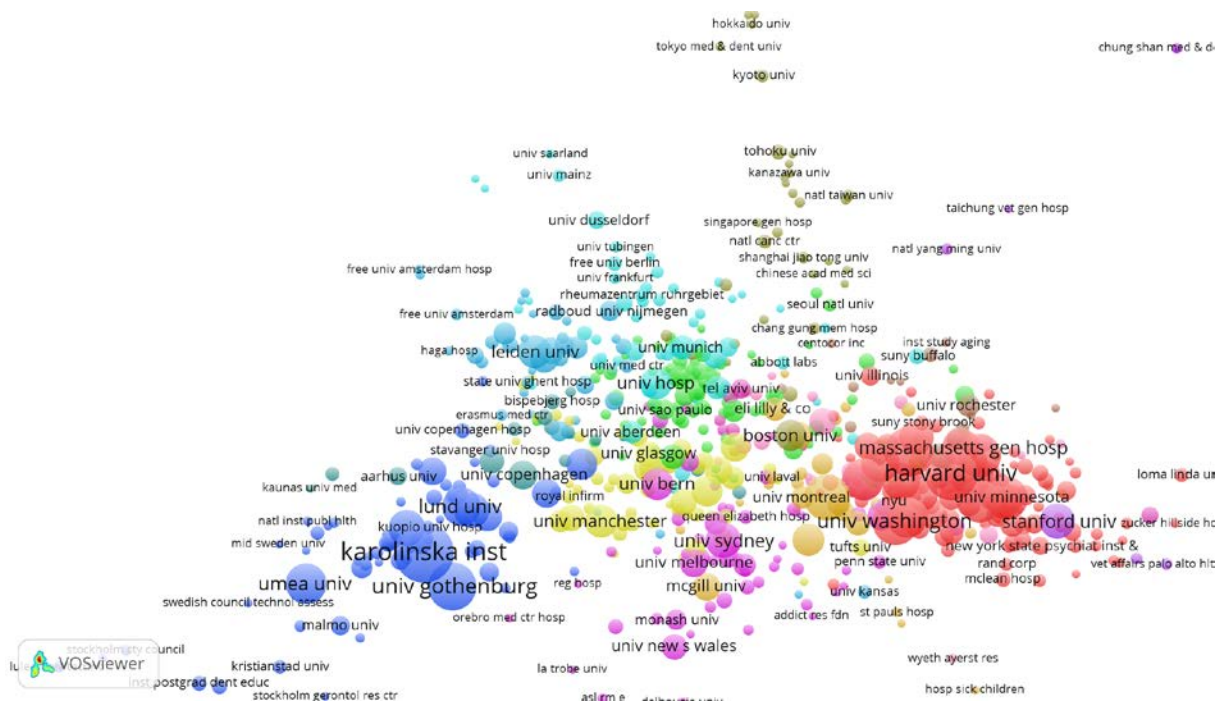
Reference age was noted by calculating the difference between clinical guideline publication year and cited reference year. The mean age of references was 7.8 years, while the maximum age was 57 years before publication of the clinical guideline.

Figure 2. Reference age up to 20 years. A total of 406 references were older than 20 years (max= 57).

Lastly, the institution of origin for the authors cited in the clinical guidelines was noted by identifying institutions from the addresses in the cited papers in WoS data. Swedish institution

names have been harmonized in the visualisation in Figure 3. It is notable that Swedish institutions are found prominently in the guidelines (mostly in the blue cluster), while U.S institutions (red) are expectedly found to a high degree. Cited literature is predominantly Western, while a small set of Asian institutions (predominantly in Japan and China) are found in the centre-top-right part of the map. These aspects of geographical features will be assessed and evaluated further in future research. It is especially relevant to compare qualitative aspects of citation practices in the clinical guidelines to traditional patterns of citation in academic literature to identify differences and biases in the kind of literature.

Figure 3. Co-author map of institutions. Of 6.631 institutions identified, authors from 665 institutions have published at least 5 papers shown here.



DISCUSSION

In this ongoing research, it is argued that it is important to widen the view of the impact of research so that impact can be measured outside the strict academic literature. The reason for this is that there is often long distance between basic academic research and the professional practice within which it is used. Therefore, clinical research must be measured with indicators that are found between the abstract research and its use in the clinic.

This project is positioned on the boundary between traditional bibliometric analysis (that typically measures intra-scientific relevance) and evidence-based *practice*, where criteria of good healthcare builds on more aspects than meta analyses of academic research.

The specific relevance of this research is related to the model for distributing national funding for the participation of Swedish county councils (the level at which healthcare is distributed) to medical education and clinical research at university hospitals as opposed to the medical schools that they are related to. These funds, the so called ALF-funds amounts to the level of

1.500 full time equivalents and between 3.000 and 3.500 publications, according to the Swedish Research Council (2013).

The goal of this project is to develop a conceptual and theoretical contribution to the development of indicators for measuring the impact of research outside of the specifically academic literature.

Outstanding issues that are pursued presently regards:

- The relationship between professional/clinical impact on the one hand and academic impact on the other. Should both kinds of impact be measured and do they complement each other or are they collapsible into each other?
- Citation window length. According to preliminary analyses, impact in clinical guidelines has a mean estimated turnover of about eight years, which is double the quite short citation window of four years that is generally used for academic impact.
- How to handle references to non-clinical research. A certain share of cited references are directed towards other clinical guidelines and reviews instead of to the underlying clinical research.
- The handling of listed, but not used references in the clinical guidelines. In many guidelines, the expert panel list not only the studies used to provide the guidelines, but also references that were judged to be of not sufficient quality. In the collection process, these references must be discarded so as not to include abandoned literature.
- Coverage for sources in PubMed amounts to almost 75 %. What other sources could be included to increase coverage and what is enough?

In sum, it is argued professional impact, here designated as impact of clinical research within clinical guidelines issued at the national or international level, provides grounds for measuring impact closer to the professional expertise that actually uses the research. Professional impact, as opposed to intra-scientific impact within the academic literature and “social impact”, based on mentions in social media or inclusion in on-line reference databases should be viewed as a distinct instance of research impact that operates on its own specific premises.

REFERENCES

- Andersen, J. P. (2013). *Conceptualising research quality in medicine for evaluative bibliometrics*. Københavns Universitet: Det Humanistiske Fakultet.
- Bornmann, L. (2012). Measuring the societal impact of research. *EMBO reports*, 13(8), 673-676.
- Bornmann, L., Haunschild, R., & Marx, W. (2015). *Policy documents as sources for measuring societal impact: How is climate change research perceived in policy documents?* Unpublished manuscript, (arXiv:1512.07071).
- Hicks, D., Wouters, P., Waltman, L., Rijcke, S. d., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature*, 520, 429-431.
- Lewis, G. (2004). Citations to papers from other documents. In H. F. Moed, W. Glänzel & U. Schmoch (Eds.), *Handbook of quantitative science and technology research: the use of publication and patent statistics in studies of S&T systems* (pp. 457-472). Dordrecht: Kluwer Academic Publishers.

- Lewison, G., & Sullivan, R. (2008). The impact of cancer research: how publications influence UK cancer clinical guidelines. *British Journal of Cancer*, 98(12), 1944-1950. doi: 10.1038/sj.bjc.6604405
- Swedish Research Council. (2013). *En utvecklad modell för kvalitetsutvärdering av klinisk forskning finansierad av ALF-medel*. Stockholm: Swedish Research Council.
- Swedish Research Council. (2015). Research quality evaluation in Sweden – FOKUS: Report of a government commission regarding a model for resource allocation to universities and university colleges involving peer review of the quality and relevance of research (pp. 123). Stockholm: Swedish Research Council.
- Thelwall, M., & Maflahi, N. (2016). Guideline references and academic citations as evidence of the clinical value of health research. *Journal of the Association for Information Science and Technology*, 67(4), 960-966.
- Wilsdon, J., et al.,. (2015). The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management. doi: 10.13140/RG.2.1.4929.1363



Technology Push / Market Pull Indicators in Healthcare

Irina Efimenko, Vladimir Khoroshevsky**, Ed Noyons***, Evgeny Nochevkin****

* Semantic Hub; National Research University Higher School of Economics, Moscow, Russia, veassi@mail.ru

** Semantic Hub; Dorodnitsyn Computing Center, Moscow, Russia, v.khor@mail.ru

***Centre for Science and Technology Studies, Leiden University, Leiden, Netherlands,
noyons@cwts.leidenuniv.nl

****Mail.Ru Group, Moscow, Russia, paperny@corp.mail.ru

ABSTRACT

Healthcare and life sciences are among the most important drivers which form the present-day landscape of science and technology in general. A whole range of emerging areas of research and disruptive technologies are related to healthcare. The applied nature of such areas of research makes it important to specify indicators which describe these areas not only from R&D, but also from user need side. We analyze the content of domain-specific social media and online consulting services in healthcare with the help of semantic technologies in order to extract widespread and emerging user needs. We will map the corresponding topics on the agenda of scientific papers in healthcare. Understanding the intersection of these two agendas and the coverage of user needs by science and technology activities leads us to the development of the “market pull” indicators for emerging areas of research.

Our research has three main stages. The first one is to discover the agenda of users of scientific discoveries and developments in medicine and healthcare, namely doctors and patients. The second one is to discover the agenda of researchers and developers. The last one is to compare the results and to interpret convergence or discrepancy of two agendas. By now we have completed stage one and in the course of stage two.

At the first stage, we have processed the content of the leading Russian online service for patients (<https://health.mail.ru/>). It contains over 300,000 pairs of posts (questions to doctors and doctors’ answers). All the posts are initially classified according to the branches of medicine with the tools developed by the Mail.Ru Group, who is hosting health.mail.ru service. 38 branches are provided. Our task was to develop a natural language processor and to process the full texts of posts. The NLP system was developed based on our previous results in semantic technologies and ontology-based information extraction (Efimenko et al. 2016) and on software tools developed by the Semantic Hub company (<http://www.semantic-hub.com/#!en/pu9y8>).

At the stage of ontology specification, we identified the most important types of posts based on patients’ questions. These types refer to the most common user needs and reasons for using online service instead of visiting a doctor. User needs include requests for diagnostics or dispatching, treatment prescription, risk forecasts, test types, etc. Reasons for using telehealth tools include lack of time for visiting a hospital, distrust for local doctors and some others. Other types refer to cross-domain topics, such as healthy lifestyle, pregnancy, etc. A post can

correspond to multiple types. A total of 65 post types were identified. User needs, reasons for using online service and the most important objects of interest were specified in the ontology. Objects of interest include medical specialties, diseases (medical conditions), symptoms, drugs, parts of the human body (incl. tissues, systems, etc.), diagnostic procedures (medical imaging technique, in vitro diagnostics) and treatment methods. Ontology-driven text processing was performed. Information extraction results were analyzed in the light of the branches of medicine represented in the content. Semantic profiles were built for each branch. These profiles specify agenda of patients (or their families and friends) and consultants. VOSviewer tools were used for visualization. An example for oncology is given in Figure 1.

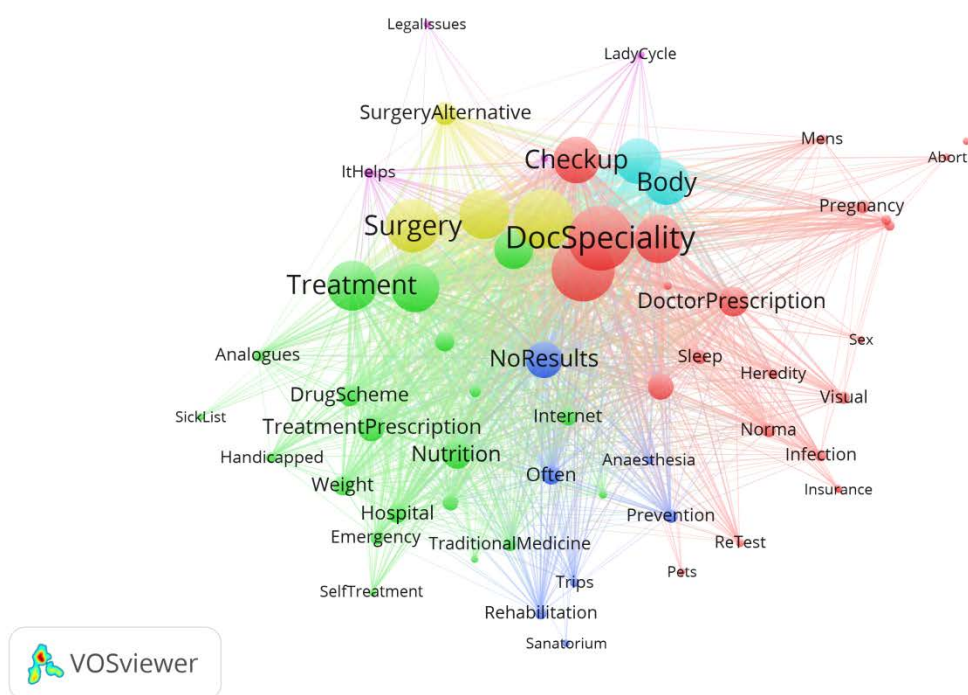


Fig. 1. Semantic profile of the oncology

The results show that some of the problems which patients face are related to the (dis-) functioning of the health system and not directly to S&T issues. Other topics and objects of interest are, however, clearly within the scope of STI problems.

The next step in our analysis was to drill down to see the most popular instances of those classes which were recognized as related to S&T issues. The examples are given in Table 1.

Table 1. Objects of interest in user posts¹

<i>Nº</i>	<i>Ontology class</i>	<i>Label on the map</i>	<i>Number of unique instances in the collection (in the field of oncology)</i>	<i>Top-5 instances (in the field of oncology)</i>
1.	Medical specialty	DocSpeciality	65	Oncologist, surgeon, pediatric surgeon, general practitioner, gynaecologist
2.	Disease (Medical condition)	Diagnosis	554	Breast cancer, lung cancer, adenocarcinoma, melanoma, stomach cancer
3.	Symptom	Symptoms	448	Tumour (various forms), pain (various types), changes in lymph nodes, temperature (fever), loss of appetite
4.	Drug (specific one or a type)	Drugs	949	<i>Types:</i> anaesthetics and analgesics, antibiotics, vitamins, cancer vaccines; <i>Specific drugs:</i> Tamoxifen, Tramadolom, Cisplatin, Xeloda (Capecitabine), Carboplatin
5.	Part of the human body	Body	350	Lymph nodes, breast, liver, abdominal cavity, lung
6.	Diagnostic procedure	Checkup, Test	173	<i>Medical imaging technique:</i> ultrasound, computer assisted tomography, magnetic resonance imaging, X-ray, colonoscopy; <i>In vitro diagnostics:</i> blood tests (various types), biopsy (histology), FNAB, biopsy (cytology), urinalysis
7.	Treatment method	Treatment	113	Chemotherapy, actinotherapy, glandular therapy, cytokine therapy, immunotherapy

In the course of analysis, we have identified new classes which should be analyzed in the drill-down mode. They include types of surgery, food and biomarkers. We will take instances of such classes into consideration in the next stages of our research.

At the second stage, the task was to process research papers. The scientific research was represented by the relevant parts of the bibliometric landscape of all sciences as defined by the CWTS publication classification scheme (Waltman & van Eck, 2012). Figure 2 shows the results for oncology based on an overview of all sciences (around 800 clusters of publications) as in Web of Science.

¹ In all the cases, objects of interest include not only oncology-specific instances, but all the objects mentioned by users (for example, heart conditions and allergies as instances of the class “Disease, Medical condition”).

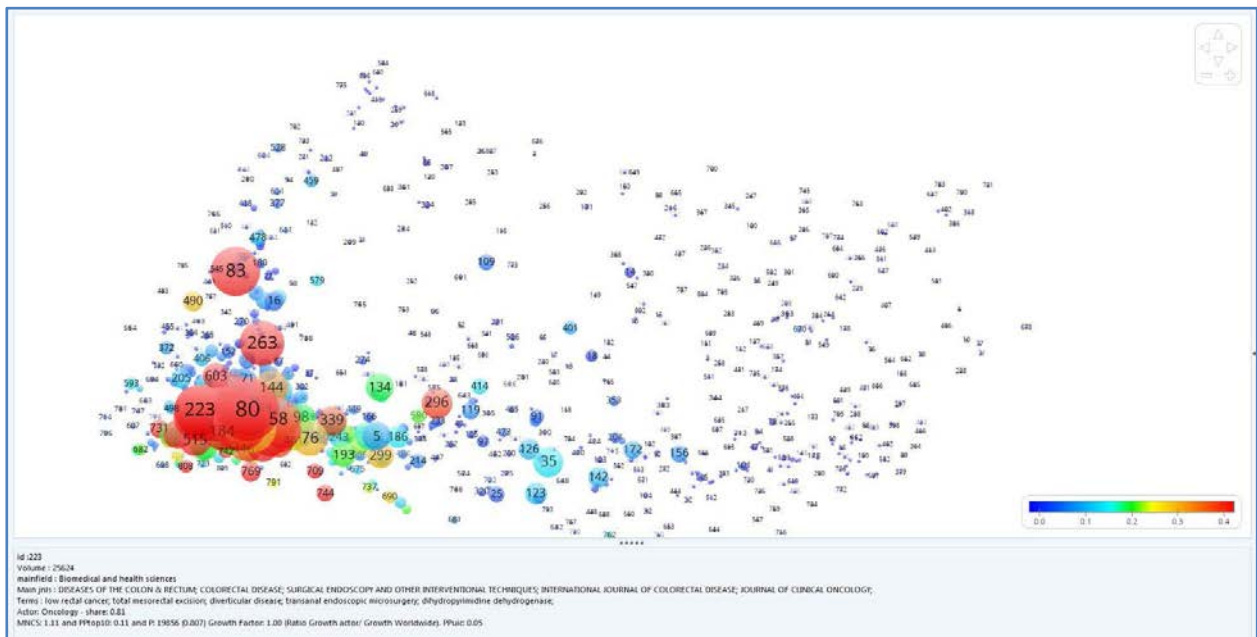


Fig. 2. Oncology in research agenda

Semantics of each cluster is represented by terms. Our next task is to “ontologize” these terms. The corresponding ontology should include the same classes as the one for user posts (medical conditions, drugs, diagnostic procedures, types of surgery, treatment methods, etc.), as well as other types relevant for research agenda. Examples are provided in Figure 3.

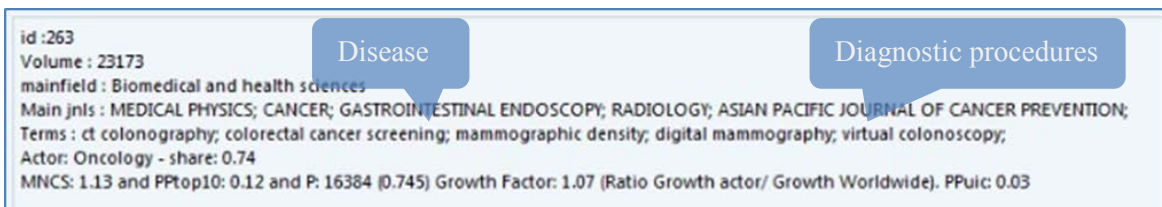


Fig. 3. Terms and ontology classes in research agenda

After we develop the ontology, we will build a representation of R&D agenda which can then be compared to users’ agenda.

Working hypotheses are the following ones:

- 1) If an important part of the users’ agenda is not covered by the research, it can be explained by disfunctioning of the health system (for non-STI problems) or it is a market pull effect and a window of opportunity for developers.
- 2) If an important part of R&D agenda does not appear in user posts, it could be interpreted as a window of opportunity for market players. In some cases, it can be explained by barriers (for example, a new diagnostic procedure or a treatment method is too expensive, unknown to practitioners, or there are unsolved legal problems).

We plan to build profiles for all the branches and to compare the degree of convergence / discrepancy of two agendas for all of them.

The most ambitious task is to develop metrics for comparing two agendas and to propose market pull indicators.

REFERENCES

- Efimenko I., Khoroshevsky V., Noyons E. (2016) Anticipating Future Pathways of Science, Technologies & Innovations: (Map of Science)² Approach. Chapter in Book: Anticipating Future Innovation Pathways through Large Data Analytics (eds.: T. Daim, A. Porter, D. Chiavetta, O. Saritas), Springer Verlag, 26 p. (in press)
- Waltman, L., & Van Eck, N.J. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology*, 63(12), 2378-2392. doi:10.1002/asi.22748.



Mapping the networks of cancer research in Portugal: first results¹

Oriana Rainho Brás*, Jean-Philippe Cointet**, João Arriscado Nunes***, Leonor David****, and Alberto Cambrosio*****

*orianarb@socius.iseg.ulisboa.pt

SOCIUS – Research Centre in Economic and Organizational Sociology, CSG, School of Economics and Management, University of Lisbon, Rua Miguel Lupi, Lisboa, 1249-078 (Portugal); ASPIC Portuguese Association for Cancer Research, Rua Dr. Roberto Frias, Porto, 4200-465 (Portugal).

** jean-philippe.cointet@polytechnique.edu

LISIS – Laboratoire Interdisciplinaire Sciences Innovations Sociétés, INRA – Institut National de la Recherche Agronomique, 5 boulevard Descartes, Champs sur Marne, Marne-la-Vallée, 77454 (France).

*** jan@ces.uc.pt

CES – Centre for Social Studies, University of Coimbra, Colégio São Jerónimo, Coimbra, 3000-995 (Portugal); Faculty of Economics, University of Coimbra, Avenida Dias da Silva, Coimbra, 3004-512 (Portugal).

****ldavid@ipatimup.pt

IPATIMUP – Institute of Molecular Pathology and Immunology of the University of Porto, Rua Júlio Amaral de Carvalho, Porto, 4200-135 (Portugal); Medical Faculty, University of Porto, Alameda Prof. Hernâni Monteiro, Porto, 4200-319 (Portugal).

*****alberto.cambrosio@mcgill.ca

Department of Social Studies of Medicine, McGill University, Peel 3647, Montreal, Quebec (Canada).

ABSTRACT

Social studies of cancer research at the international level have contributed to a better understanding of the developmental dynamics – both organizational and epistemic – of this field (Keating & Cambrosio, 2012). In contrast, despite its robust development, oncology research in Portugal has been the subject of only few studies. Most of them have a strong focus on the first half of the 20th century (Raposo, 2004; Costa, 2010, 2012a; 2012b), while a few focus on more contemporary events (Nunes, 2001).

Consequently, we do not have a clear picture of recent trends in oncology research in Portugal, and how it integrates into the international landscape. This hinders public accountability of oncology research while also limiting the analysis of how this research relates to health care delivery, health outcomes, and health policy formulations.

This paper presents the first results of an ongoing research project on the organizational and epistemic development of oncology research in Portugal, covering the period from the end of the 20th century to 2015. Among other issues, we intend to explore the extent to which oncology research in Portugal mirrors the international dynamics at a smaller scale, and the extent to which it presents features of its own.

The study draws upon computer-based analysis of publications using the platform CorText (<http://www.cortext.net/>) of IFRIS (Institut Francilien Recherche, Innovation, Société), along with interviews with Portuguese oncologists and related practitioners.

¹ This work was supported by by Fundação Calouste Gulbenkian, Portugal, and ASPIC - Portuguese Association for Cancer Research, Portugal

The database of publications is derived from queries in Web of Knowledge and PubMed. We searched Web of Science for publications whose subject was oncology and whose country was Portugal. Alongside, we searched PubMed for publications that were indexed with the MeSH term “Neoplasms” and presented Portugal as their affiliation. PubMed does not include information on cited references, and only recently includes the addresses of all co-authors. To be able to include this information we searched for the publications found in PubMed in Web of Science.

This process retrieved 5725 publications dated between 1976 and 2015, with a marked increase in the number of publications beginning at the turn of the century. Deep changes in the Portuguese political context partially account for this growth. In 1995, the government created an autonomous Ministry of Science, whereas prior to that time research was part of the Education Ministry. In 1997 the Foundation for Science and Technology was created within the Ministry of Science, in charge of developing, evaluating and funding scientific research. The creation of scientific research institutes – autonomous from universities but affiliated, or closely collaborating with them – contributed to the development of research in general and of the health sciences in particular, since some of those institutes were working in that field. But this is only part of the story. Portuguese oncology research also developed in close connection with changes at the international level that can be broadly characterized as a molecular turn in oncology. These changes include: (a) the adoption in the 1990s of molecular pathology methods that made possible the stratification of patients on the basis of molecular biomarkers, and the related reconfiguration of clinical trials and of our understanding of different types of cancer as consisting in fact of multiple diseases; (b) the emergence at the turn of the century of targeted therapies; and (c) the concurrent adoption of high-throughput methods, first microarrays, then next-generation sequencing.

We looked at institutional collaboration as shown by co-authorships. Starting from the raw co-occurrence matrix featuring the joint number of papers published by two institutions, only links between two institutions whose shared number of papers was more numerous than expected from a null model were conserved. More precisely, we computed Cramer distance between any couple of institutions and pruned the network under a certain value of proximity. As an outcome we have two collaboration networks. One is made of every collaboration event between institutions, which features links between two institutions whose strength is directly proportional to their shared number of papers. The second network – which we will call specific collaboration network – highlights links which are statistically surprising in the raw collaboration network.

When considering every collaboration, we see a star-like map with a few institutions at the centre. In the specific collaboration network, we see regional clusters of collaboration also structuring this field in Portugal. There are also significant collaborations between Portuguese institutions and European countries, Brazil and the USA. A more thorough exploration of this data will show us whether, and to what extent, these collaborations stem from a considerable number of Portuguese researchers leaving the country to obtain their PhD and/or accept positions in foreign institutions. Moreover, we will further investigate whether, as hinted by maps of the most recent period, we can differentiate between institutions that privilege regional and national collaborations, and a few others whose strategy is to rely on international relations.

Oncology is a multidisciplinary field. We examine it without defining pre-established boundaries about who or what should count as relevant actors and contexts. Increasingly, research in oncology is defined by the presence of heterogeneous collectives and networks of human and non-human actors, *i.e.* bio-clinical collectives of medical oncologists,

pathologists, biostatisticians, and researchers in molecular oncology and cancer genomics, as well as research tools, biological entities, pathogenic phenomena, and clinical techniques. Attesting to the diversity of relevant actors in oncology research in Portugal is the large number of publications in a specialized journal on the psychological aspects of cancer. Most of those publications are meeting abstracts. While relevant in terms of the number of publications, this subfield is absent from maps displaying journal co-citation and reference co-citation networks (computed for the top 50 journals and top 50 cited references), and relatively marginal in inter-citation maps (computed for top 200 journals and cited journals), *i.e.* it is not significantly connected to other subfields in oncology research. This feature is arguably present in countries other than Portugal. It can indicate an interesting example of an innovation coming from the outside of the core components of a research field, a hypothesis we will further explore in this study.

Natural Language Processing tools can be used to examine semantic networks of terms used in titles, abstracts and keywords of the retrieved publications, and thus to analyse the main themes characterizing oncology research in Portugal. Initial semantic maps show that research focuses on gastric cancer, breast cancer, thyroid cancer, prostate cancer, lymphoma, leukaemia, prognosis, metastasis, cancer epidemiology, and cancer cell biology. Further research will fine-tune this analysis by examining the evolution of these subfields and of their connections during the last 25 years. Semantic maps also display the multiple, heterogeneous assemblages that compose the networks of oncology knowledge production, which include tools, concepts, methods, patients, families, micro-organisms, biological processes, genes, cells, drugs, and tumours.

Connecting the semantic maps with the research institutions and their degree of specialization in the various themes over time will allow us to examine the evolution of the political, organizational, and epistemic strategies of these institutions, as well as the consequences and implications of these strategic choices.

REFERENCES

- Costa, R.M.P. (2010). Discurso médico, saúde pública e estratégias políticas para “Uma questão palpitante do tempo actual”: a emergência da luta contra o cancro em Portugal (1904-1923). *Revista da Faculdade de Letras. HISTÓRIA*, 11(1), 135-164.
- Costa, R.M.P. (2012a). Between parasitic theory and experimental oncology: a proposal for systematizing oncological science in Portugal, 1889-1945. *História, Ciências, Saúde – Manguinhos*, 19(2), 409-430.
- Costa, R.M.P. (2012b). Instituto Português de Oncologia: organização e dinâmica assistencial (1928-1958). *Estudos do Século XX*, 12, 311-333.
- Keating, P. & Cambrosio, A. (2012). *Cancer on trial: oncology as a new style of practice*. Chicago: University of Chicago Press.
- Nunes, J.A. (2001). Laboratórios, escalas e mediações na investigação biomédica. A oncobiologia entre o global e o local. In J.A. Nunes & M.E. Gonçalves (Orgs.), *Enteados de Galileu: a semi-periferia no sistema mundial da ciência* (pp. 33-75). Porto: Afrontamento.
- Raposo, H. (2004). A luta contra o cancro em Portugal. Análise do processo de institucionalização do Instituto Português de Oncologia. *Fórum Sociológico*, 11/12 (2ª série), 177-203.



Access to global health research. Prevalence and cost of gold and hybrid open access¹

Stefanie Haustein^{*}, Elise Smith^{*}, Philippe Mongeon^{*}, Fei Shu^{**} and Vincent Larivière^{*,***}

^{*}stefanie.haustein@umontreal.ca; elise.smith@umontreal.ca; philippe.mongeon@umontreal.ca
École de bibliothéconomie et des sciences de l'information, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal (Canada)

^{**}fei.shu@mail.mcgill.ca
School of Information Studies, McGill University, 3661 Peel, Montréal (Canada)

^{***}vincent.lariviere@umontreal.ca
Observatoire des Sciences et des Technologies (OST), Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Université du Québec à Montréal, Montréal (Canada)

ABSTRACT

As it is a priority of global health research (GHR) to achieve equity in health worldwide, there is an increased demand and expectation that knowledge be shared freely and without barriers. Making research findings available for free to readers by publishing open access (OA) is thus central to GHR.

Several studies have assessed the extent to which different forms of OA prevail but despite the importance of free access to knowledge in GHR, particular empirical evidence is missing. This paper aims to fill this gap by analyzing the extent to which GHR papers indexed in PubMed are published OA and how much it costs to publish in gold and hybrid OA journals. Findings show that between 2010 and 2014 as few as 18% of papers were published in gold OA journals, 7% published as hybrid OA (i.e., OA papers in subscription journals), while more than 60% were behind paywalls. Costs for gold OA amounted to \$990,619 for 404 papers, whereas \$722,631 were spent on article processing charges (APCs) of 223 hybrid papers. The majority of APCs were obtained by large commercial publishing houses known for exorbitant profit margins.

INTRODUCTION

The field of global health can be defined as an "area for study, research, and practice that places a priority on improving health and achieving equity in health for all people worldwide" (Koplan et al., 2009, p. 1995). To achieve this goal, global health research (GHR) partnerships or collaborations often include a myriad of players including researchers from various countries, key stakeholders, non-governmental organizations, governmental institutions and for-profit organisations (Szlezák et al., 2010). These players share knowledge and expertise to solve complex health issues of global importance. In such a context, there is an increased demand and expectation that knowledge be shared with all actors, promoting its use in knowledge translation and subsequent research. A commonly used method to provide improved access is to publish research findings open access (OA), making it free to all readers. However, substantial and ever increasing subscription fees of journals pose a cost barrier to individuals and organisations with limited resources. The use of OA publication in

¹ This work was supported by the Canada Research Chair on the Transformations of Scholarly Communication.

GHR has been promoted by various scholars and in policy initiatives as a method of facilitating knowledge transfer (Siriwardhana, 2015), and it is seen as a necessary method to "building science capacity in developing countries" (Chan, Kirsop, & Arunachalam, 2005, para. 1). Certain programs are specifically set up to provide institutions in low and medium income countries (LMICs) access to biomedical and health literature at no or very low cost. For example, through HINARI² publishers such as Wiley-Blackwell, Reed-Elsevier, Wolters Kluwer and Springer Nature provide access to 1,500 journals. Knowledge users in eligible countries can, however, only access the content if they are affiliated with a registered institution, which represents a restriction to access for many actors working in GHR.

Several studies have assessed the extent to which different forms of OA prevail (e.g., Björk, 2012; Björk et al., 2010; Dallmeier-Tiessen et al., 2010; Gargouri et al., 2010; Laakso et al., 2011; Laakso & Björk, 2012). These studies usually distinguish between green (i.e., self-archived pre- or post-prints) and gold OA papers. The latter can be further broken down into papers published in gold OA journals with and without article processing charges (APCs), delayed OA journals or hybrid journals (i.e., subscription journals offering article-based OA options in exchange for an APC). Björk et al. (2010) estimate that irrespective of the discipline 20.4% of papers published in 2008 were available online for free (8.5% gold or hybrid OA, and 11.9% Green OA). When comparing different disciplines of research, gold OA is particularly prevalent in medical journals (13.9%) (Björk et al., 2010). The most recent and largest OA study estimates that 12.1% of papers published between 2011 and 2013 appeared in gold or hybrid OA journals and 46.9% of papers were freely available online combining any type of unrestricted access (e.g., institutional repositories, personal websites) (Archambault et al., 2014). For the field of Public Health & Health Services, which represents the most comparable to GHR, Archambault and colleagues (2014) report even higher rates of overall OA (57.2%) and gold or hybrid OA (15.8%) rates.

Despite the importance of free access to GHR, there is no empirical evidence on the uptake of OA in the field. This paper aims to fill this gap by answering the following research questions:

1. What types of publication practices are prevailing in GHR?
 - a. To what extent do journals allow gold, hybrid or green OA?
 - b. To what extent do authors make use of various routes of OA?
2. What are the costs of gold and hybrid OA in GHR?
 - a. What are the average prices of gold and hybrid APCs?
 - b. Which publishers benefit most from gold and hybrid APCs?

DATA AND METHODS

Defining relevant papers

GHR papers were identified searching PubMed for documents indexed with the Medical Subject Heading (MeSH) term "Global Health" published between 2010 and 2014³. The result of 4,333 documents was restricted to document types reporting research results such as articles and reviews, excluding document types such as editorials and news items. The 3,461

² <http://www.who.int/hinari/en/>

³ Query "Global Health"[MeSH Major Topic] AND 2010[PDAT] : 2014[PDAT] carried out on PubMed on 16 February 2016.

remaining documents were downloaded from PubMed, 95 documents were further excluded as their publication year was outside the selected 5-year period or they turned out to be book chapters. The resulting data set includes 3,366 GHR journal articles, which were published in 909 journals.

Determining access status for papers, journals and publishers

Two types of OA can be identified for papers on Pubmed: “free article” or “free PMC article”. The former signifies gold or hybrid articles (i.e., articles that are freely available on the publisher’s site) and the latter indicates deposits in PubMed Central (PMC) made by publishers (often after an embargo period) or by authors. However, several gold OA journals deposit directly in PMC, in which case “free in PMC” indicates gold OA articles. In this study, OA was defined as Gratis not Libre OA and thus entails access to the content free of charge but possibly involving copyright and licensing restrictions (Suber, 2008). The journal status (gold, hybrid or subscription only) was initially determined using data from the Directory of Open Access Journals (DOAJ), Ulrich’s Periodicals Directory and journal lists from Reed-Elsevier, Sage, Springer, Taylor & Francis and Wiley-Blackwell. Since there was conflicting information between different data sources, it was decided to manually verify publication policies for all journals, classifying them in the following categories:

- **Gold OA journal (non-APC):** Open access journal which provides immediate access to all of their content free of charge to both readers and authors. Many of these journals were often financed or subsidized by scientific societies or associations.
- **Gold OA journal (APC):** Open access journal which provides free immediate access to all of their contents based on an author-pays model via APCs. APCs were collected in or converted to USD⁴.
- **Delayed OA journal:** Subscription journal which provides all content for free after an embargo or delay period of several months to years. Journals which provide delayed open access to only some of their content were classified as subscription journals and their free papers identified as delayed OA articles.
- **Hybrid journal:** Subscription journal which is primarily financed by reader-pays model based on subscriptions and pay-per-view fees but allows authors to pay an APC to make their article available free of charge for the reader without delay. APCs were collected in or converted to USD.
- **Subscription only journal:** Subscription journal which is financed by reader-pays model based on subscriptions and pay-per-view fees and does not offer author-pays OA options. Some subscription journals might decide to make single articles available for free temporarily or permanently to promote certain contents. If identified by PubMed, these free articles are coded as “other free access”.
- **Unknown:** Journals for which the access status could not be determined.

Combining PubMed’s article level information with the journal status, papers were classified as follows:

- **Gold OA article (non-APC):** “free article” and “free PMC article” published in gold OA journal (non-APC).

⁴ If APCs were not provided in USD, currencies were converted using the mean of weekly historical conversion rates between 01.01.2010 and 31.12.2014 using OANDA (<http://www.oanda.com/currency/historical-rates>).

- **Gold OA article (APC):** “free article” and “free PMC article” published in gold OA journal (APC).
- **Delayed OA article:** “free article” published in delayed OA or subscription journal with delayed OA option.
- **Hybrid article:** “free article” published in hybrid OA journal.
- **Toll access:** “not free” article published in any type of journals.
- **Green OA article (PMC):** “free PMC article” published in hybrid OA or subscription journal.
- **Other free access:** “free article” in subscription journal or journal for which status is unknown.

Publisher affiliations were determined based on the above-listed sources and the Web of Science (WoS) and manual searches in case of conflicting information. Acquisitions and mergers of journals and publishing houses were accounted for by assigning journals to the most recent publisher using the method described in Larivière, Haustein, and Mongeon (2015).

Limitations

The analysis of green OA is limited to PMC deposits only, thus ignoring other forms on self-archiving in institutional repositories or on authors’ websites or academic social networks. The journal status and publisher affiliation was determined based on information provided on the journal website at the time of analysis. Publishers, journal policies and APCs might thus have been different at the time of publication.

Publishers of hybrid and subscription journals sometimes provide discounts for OA publication particularly to authors from LMICs which would change or cancel the APC. However, it remains impossible to identify the publications that have been published at discounted prices and thus it was assumed that the full APC was paid.

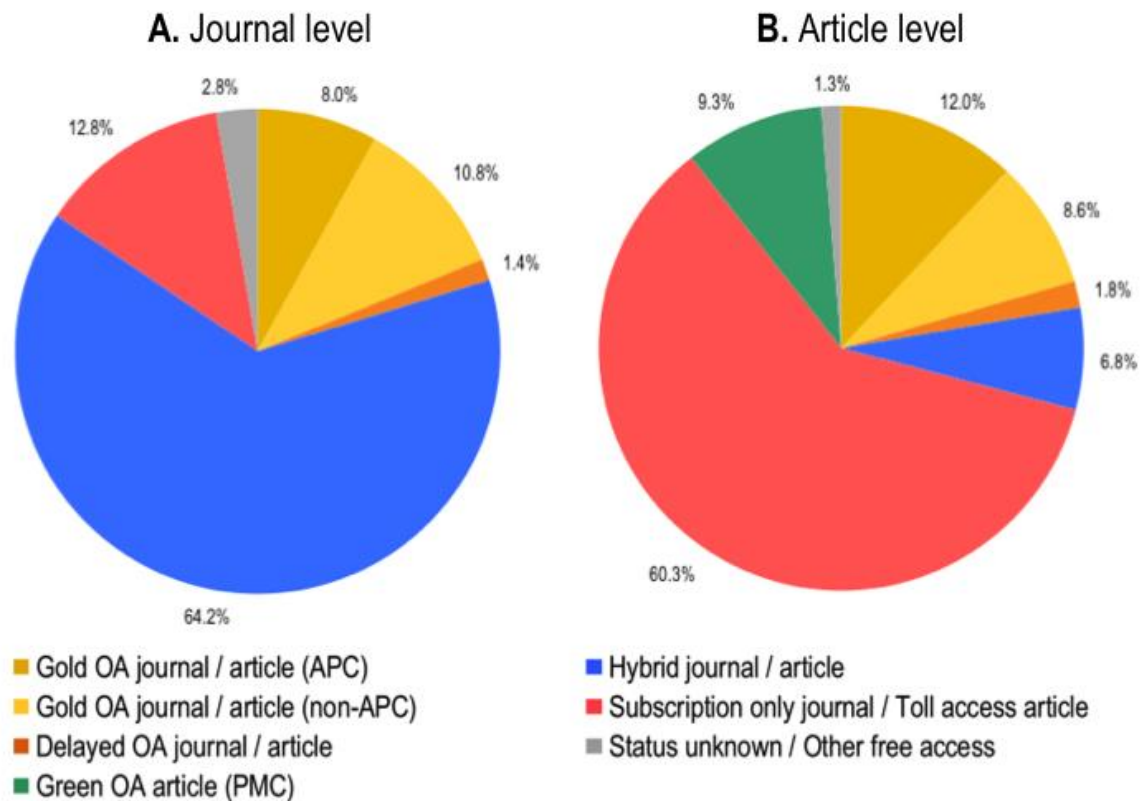
We rely on the paper status provided by PubMed and might thus mistake temporary free access for hybrid articles or disregard this form of transient OA altogether (Archambault et al., 2014). To ensure the reliability of the paper level OA status provided by PubMed, we manually verified 235 papers, for which the OA status contradicted the recorded journal policy. In 114 cases, the information was corrected changing the access status from “not free” to “free article”.

RESULTS AND DISCUSSION

The majority of journals publishing GHR papers were hybrid journals (64.2%), which is mostly due to major publishers such as Reed-Elsevier and Wiley-Blackwell, who offer hybrid options for most of their journals. Around one fifth of journals were gold OA—more than half of which were actually not charging any publication fees—while 12.8% were subscription journals without open access options (Figure 1A). Prominent Gold OA journals without APCs included the *Bulletin of the World Health Organization* (38 GHR papers 2010-2014), *Morbidity and Mortality Weekly Report* (32) and *Weekly Epidemiological Record* (20), while *Lancet Global Health* (54), *PLoS Medicine* (53), *Global Health* (49), *Global Health Action* (36), *BMC Public Health* (24) and *PLoS One* (20) were publishing the largest number of GHR papers among gold OA journals with APCs. In the context of GHR, *Lancet* (315), *Global*

Public Health (55), *BMJ* (31) and *Social Science & Medicine* (30) were the most productive hybrid journals, while *Revue scientifique et technique* (49), *New England Journal of Medicine* (39), *JAMA* (26) and *Nature* (19) published the most papers among subscription journals without an author-paid OA option. This demonstrates the popularity of publishing in prestigious high-impact journals despite the lack of gold or hybrid OA.

At the article level it becomes apparent that the majority of authors do not make use of gold or hybrid OA (Figure 1B), as more than half of all articles (60.3%) are locked behind a paywall. Combining gold and hybrid access, more than one quarter (27.4%) of articles are available on the publisher's website at no charge for the reader and without a delay, which exceeds the 15.8% reported for Public Health & Health Services by Archambault et al. (2014). Among gold OA options, publishing in journals with an APC (12.0% of papers) is more popular than those without (8.6%), which might, at least to a certain extent, be caused by differences in journal reach, scope and quality. While hybrid OA was available for 2,102 papers, authors of 229 opted to pay the APC, which represents an uptake of 10.9%. This represents a particularly high rate compared to Björk (2012) who, based on the low uptake of hybrid OA, concluded that it was a failed experiment by publishers. Hybrid uptake differs between journals and publishers, which might be a result of both different APC rates and successful marketing strategies. For example, among the publishers offering hybrid for more than 35 papers in the GHR data set, hybrid uptake was 82.1% for Oxford University Press, 18.5% for Wiley-Blackwell, 13.8% for Wolters Kluwer, 6.1% for Taylor & Francis, 3.5% for Reed-Elsevier, 2.2% for the BMJ Group, 1.2% for Springer Nature, while no paper published in the GHR set used the hybrid option with Sage. The full text of 9.3% papers published in subscription or hybrid journals were available on PMC. Although this represents just one specific form of green OA for the GHR papers under analysis, it slightly exceeds the rate of 7.8% green OA for medical papers estimated by Björk et al. (2010).

Figure 1: Percentage of journals (A) and articles (B) per type of access category.

The total cost to publish OA GHR papers amounts to \$990,619 for 404 gold OA papers and \$722,631 for 223 hybrid papers⁵. Rates varied substantially per journal from \$10 (*Saudi Medical Journal*) to \$5,000 (*Lancet Global Health*) for gold OA journals and from \$774 (*The New Zealand Medical Journal*) to \$5,000 (12 journals including *Cell* and several *Lancet* journals) for hybrid journals. The presence of large for-profit publishing houses and hybrid journals among the highest APCs is striking: out of the 35 journals with APCs of \$4,000 and above, 17 are published by Wolters Kluwer, 9 by Reed-Elsevier and all but one are hybrid journals.

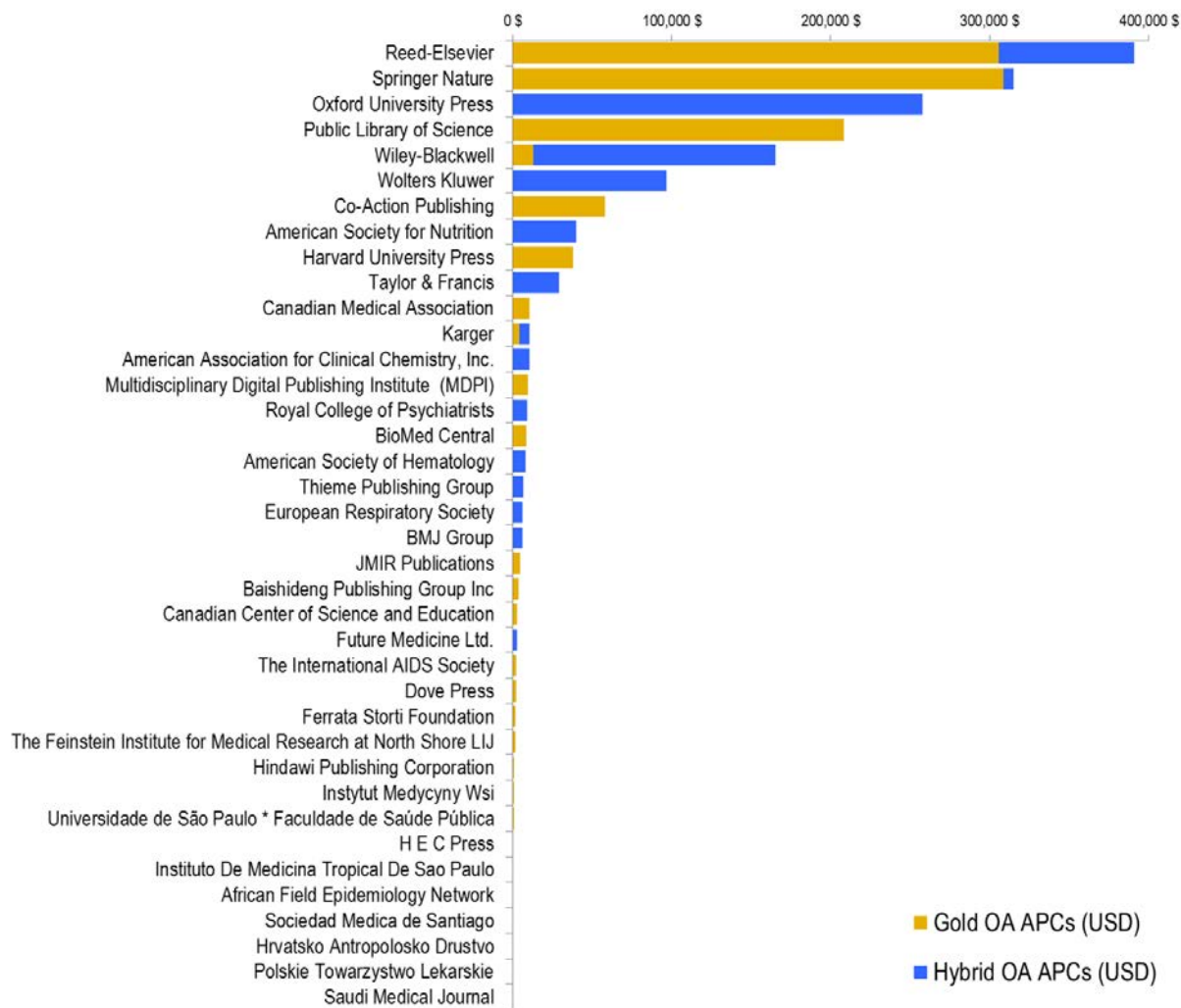
On average, APCs amounted to \$1,864 in gold OA journals, while hybrid fees were substantially higher at \$2,978. This is quite counterintuitive, as hybrid journals are already financed by subscriptions, while gold OA journals are solely based on APCs. These results corroborate findings by Solomon and Björk (2012) and van Noorden (2013) for all fields combined. Past research analyzing average APCs reported between \$660 (Outsell according to Van Noorden, 2013), \$906 (Solomon & Björk, 2012) and \$1,255 (Björk & Solomon, 2015) per paper.

Figure 2 shows the sum of APCs paid to the different publishers for gold and hybrid OA papers. The largest amounts were obtained by Reed-Elsevier (\$391,050), Springer Nature (\$307,165), Oxford University Press (\$265,695), PLOS (\$208,350), Wiley-Blackwell

⁵ The cost for the 6 hybrid papers published in the *Journal of Dental Education* were excluded from the cost calculations, as the APC could not be determined;

(\$165,300) and Wolters Kluwer (\$96,600); most of these publishers have profit margins around 30% and belong to an oligopoly or academic publishing houses based on the number of articles published (Larivière et al., 2015).

Figure 2: Sum of gold and hybrid APCs per publisher.



CONCLUSIONS AND OUTLOOK

Even if knowledge sharing is a central premise of GHR, this study shows that the majority of papers are hidden behind paywalls, and that OA rates are negligibly higher than that of other medical disciplines. Despite the importance of such knowledge to communities in the developing world, more than 60% of all papers are not freely available. Such lack of access has been recently discussed in light of the Ebola outbreak; which may have been mitigated, if scientific information had been freely available to Liberian researchers (Dahn, Mussah, & Nutt, 2015). Although hybrid OA solves the access problem, it makes the GHR community pay twice for the same content; hybrid APCs are added to the subscription fees already paid by research institutions. In fact, a significant amount of money goes into the pockets of big commercial publishers, who are known for their exorbitant profit margins.

In this study, the analysis of green OA did not extend beyond the PMC platform. Future research will examine whether the 909 journals allow self-archiving and the extent to which authors do so; this will help in understanding the broader picture of OA practices in GHR. Based on the WoS subset of papers, citation rates and citing countries will be analyzed to determine the extent to which the access status leads to different citation patterns.

REFERENCES

- Archambault, É., Amyot, D., Deschamps, P., Nicol, A., Provencher, F., Rebut, L., & Roberge, G. (2014). *Proportion of Open Access Papers Published in Peer-Reviewed Journals at the European and World Levels 1996-2013*. (RTD-B6-PP-2011-2: Study to develop a set of indicators to measure open access). European Commission. Retrieved from http://www.science-metrix.com/files/science-metrix/publications/d_1.8_sm_ec_dg-rtd_proportion_oa_1996-2013_v11p.pdf
- Björk, B.-C. (2012). The hybrid model for open access publication of scholarly articles: A failed experiment? *Journal of the American Society for Information Science and Technology*, 63(8), 1496–1504. <http://doi.org/10.1002/asi.22709>
- Björk, B.-C., & Solomon, D. (2015). Article processing charges in OA journals: relationship between price and quality. *Scientometrics*, 103(2), 373–385. <http://doi.org/10.1007/s11192-015-1556-z>
- Björk, B.-C., Welling, P., Laakso, M., Majlender, P., Hedlund, T., & Guðnason, G. (2010). Open Access to the Scientific Journal Literature: Situation 2009. *PLoS ONE*, 5(6), e11273. <http://doi.org/10.1371/journal.pone.0011273>
- Chan, L., Kirsop, B., & Arunachalam, S. (2005). Open access archiving: the fast track to building research capacity in developing countries. Retrieved from <https://tspace.library.utoronto.ca/handle/1807/4415>
- Dahn, B., Mussah, V., & Nutt, C. T. (2015, April 7). Yes, We Were Warned About Ebola. *The New York Times*. Retrieved from http://www.nytimes.com/2015/04/08/opinion/yes-we-were-warned-about-ebola.html?_r=2
- Dallmeier-Tiessen, S., Goerner, B., Darby, R., Hyppoelae, J., Igo-Kemenes, P., Kahn, D., ... van der Stelt, W. (2010). *Open Access Publishing - Models and Attributes*. Retrieved from <http://edoc.mpg.de/478647>
- Gargouri, Y., Hajjem, C., Larivière, V., Gingras, Y., Carr, L., Brody, T., & Harnad, S. (2010). Self-Selected or Mandated, Open Access Increases Citation Impact for Higher Quality Research. *PLoS ONE*, 5(10), e13636. <http://doi.org/10.1371/journal.pone.0013636>
- Koplan, J. P., Bond, T. C., Merson, M. H., Reddy, K. S., Rodriguez, M. H., Sewankambo, N. K., & Wasserheit, J. N. (2009). Towards a common definition of global health. *The Lancet*, 373(9679), 1993–1995. [http://doi.org/10.1016/S0140-6736\(09\)60332-9](http://doi.org/10.1016/S0140-6736(09)60332-9)

- Laakso, M., & Björk, B.-C. (2012). Anatomy of open access publishing: a study of longitudinal development and internal structure. *BMC Medicine*, *10*(1), 124. <http://doi.org/10.1186/1741-7015-10-124>
- Laakso, M., Welling, P., Bukvova, H., Nyman, L., Björk, B.-C., & Hedlund, T. (2011). The Development of Open Access Journal Publishing from 1993 to 2009. *PLoS ONE*, *6*(6), e20961. <http://doi.org/10.1371/journal.pone.0020961>
- Larivière, V., Haustein, S., & Mongeon, P. (2015). The Oligopoly of Academic Publishers in the Digital Era. *PLOS ONE*, *10*(6), e0127502. <http://doi.org/10.1371/journal.pone.0127502>
- Siriwardhana, C. (2015). Promotion and Reporting of Research from Resource-Limited Settings. *Infectious Diseases*, *8*, 25–29. <http://doi.org/10.4137/IDRT.S16195>
- Solomon, D. J., & Björk, B.-C. (2012). A study of open access journals using article processing charges. *Journal of the American Society for Information Science and Technology*, *63*(8), 1485–1495. <http://doi.org/10.1002/asi.22673>
- Suber, P. (2008). Gratis and libre open access. *SPARC Open Access Newsletter*, (124). Retrieved from <http://legacy.earlham.edu/~peters/fos/newsletter/08-02-08.htm>
- Szlezák, N. A., Bloom, B. R., Jamison, D. T., Keusch, G. T., Michaud, C. M., Moon, S., & Clark, W. C. (2010). The Global Health System: Actors, Norms, and Expectations in Transition. *PLOS Med*, *7*(1), e1000183. <http://doi.org/10.1371/journal.pmed.1000183>
- Van Noorden, R. (2013). Open access: The true cost of science publishing. *Nature*, *495*(7442), 426–429. <http://doi.org/10.1038/495426a>



Scientific research on diseases: the distinct profile of developed and developing countries

Alfredo Yegros-Yegros*, Ismael Rafols**, María-Francisca Abad-García***, Rogério Mugnaini****, Ingeborg Meijer*****

* a.yegros@cwts.leidenuniv.nl

Centre for Science and Technology Studies (CWTS), Leiden University, PO Box 905, 2300 AX Leiden (The Netherlands)

** i.rafols@ingenio.upv.es

Ingenio (CSIC-UPV), Universitat Politècnica de València, València (Spain)
SPRU—Science and Technology Policy Research, University of Sussex, Brighton (England)
Observatoire des Sciences et des Techniques (OST-HCERES), Paris (France)

*** abad@uv.es

Universitat de València, Departamento de Historia de la Ciencia y Documentación Avda. Blasco Ibáñez, 15. 46010 Valencia (Spain)

**** mugnaini@address.comusp.br

School of Communication and Arts, University of São Paulo, Av. Prof. Lúcio Martins Rodrigues 443, São Paulo, 05508-020 (Brazil)

***** i.meijer@cwts.leidenuniv.nl

Centre for Science and Technology Studies (CWTS), Leiden University, Leiden (The Netherlands)

INTRODUCTION

In most of the countries across the globe, health is seen as a priority. The European Commission (2013), for example, considers health as a precondition for economic prosperity given that people's health influences economic outcomes in terms of productivity, labor supply, human capital and public spending. Accordingly, the Commission places health in one of the big Societal Challenges ('health, demographic change and wellbeing') in Horizon 2020.

Part of the investment on research and development is devoted to specific diseases. In order to assess whether scientific research is targeting the most pressing diseases, some studies have tried to analyze the degree of alignment between the funding allocated to specific diseases and the burden of disease (e.g. Gillum et al, 2011; Kingel et al, 2014). Others, like Evans et al (2014), focus on the relationship between research outputs dealing with specific diseases and the burden of disease.

In the later study, Evans et al (2014) found that there was no relationship between the burden of disease and the total health research at the world level. Only when the relationship between research outputs and burden of disease was analyzed at the level of individual countries, a significant association was found between the two. Another finding of this study is the striking disparity among countries in the capacity to produce health research: developed countries publish much more biomedical research than less developed countries. While this is not a surprising finding, the authors consider that this fact combined with the tendency of focusing on national health needs, results in the overrepresentation diseases more prevalent in

developed countries and the underrepresentation of diseases affecting less developed countries.

An interesting remark by the authors of this study is that those diseases that affect most developing countries only slightly affect the developed countries (e.g. neglected tropical diseases or NTD ¹), while diseases that most afflict rich populations also affect substantially less developed countries (e.g. cancer).

Building on the study by Evans et al (2014) we put forward the following hypothesis:

H1: Developed countries concentrate most of their publications in diseases that most affect these countries while only a minor share of publications is devoted to NTD

H2: Developing countries present a more balanced publication profile, covering both NTD and also diseases that most affect developed countries

While both developed and developing countries might publish scientific articles on NTD as well as on diseases that primarily affect developed countries, we do not expect only differences in terms of the amount of publications developed and developing countries published on each type of disease, but also regarding the status of the journal in which the research was published as well as in terms of citation impact achieved, given the unequal scientific impact observed across countries (e.g. King, 2004).

H3 Regardless the type of disease, developed countries publish research in high impact journals, while developing countries publish in journals with lower impact

H4 Citation impact achieved by developed countries is higher compared to developing countries, regardless the type of research

Scientific basic research is just the first step in the development of new drugs, additional stages are needed before patients can benefit of such a drug. Preclinical research, clinical research and post-marketing are also part of the costly process of generation of new drugs. As highlighted by Wilder and Solovy (2005), probably the most important gap is that existing between basic research and pre-clinical research, as this generally requires the investment by private companies to continue the process where the public sectors left off. However, companies are more likely to invest in the development of new drugs if they can ensure a return for that investment, by introducing the new drugs in the market. Evans et al (2014) found a significant positive relationship between the market size and publication of research. The market size is not determined by the amount of people suffering from a specific disease but is related to the purchasing power of the population. Although scientific publications can be considered to be far away from final marketable drugs, we put forward the following hypothesis:

H5: Private companies are more likely to engage in public upstream research on diseases that affect primarily developed countries, both in terms of conducting and funding research

H6: Universities and Public Research Organizations lead the research on NTD, while funding to conduct this research flows mainly from Governments and NGOs

¹ http://www.who.int/neglected_diseases/diseases/en/

In the next section we describe the data and methods used in the study. Finally, we present some preliminary findings

METHODS

In order to conduct the study, we limit ourselves to a short list of diseases and countries, which will allow us to develop a more detailed analysis. Also, it is necessary to have a classification of the countries in terms of degree of development as well as a classification of diseases, especially to determine those that are prevalent both in developed and developing countries.

Selected countries

Six specific countries were selected to develop this study. The main criteria in the selection of these countries was to obtain a balanced representation of countries according to their degree of development. We used the 2015 edition of the Human Development Index (HDI)², created by the United Nations Development Programme, in order to determine the degree of development of the countries. The HDI is a composite indicator which takes into account several dimensions such as life expectancy, education, and income per capita to estimate the degree of development.

This HDI group all the countries in four categories, ranging from ‘very high human development’ to ‘low human development’. We selected countries in the three top categories of development, as countries in the lower category (‘low human development’) hardly contribute to the international scientific literature, somehow reflecting their low activity on scientific research. Table 1 shows the countries selected, representing distinct stages of development.

Table 1. Countries included in the study according to their HDI (position in the ranking)

Very high	High	Medium
Netherlands (5)	Russia (50)	Colombia (97)
Spain (26)	Brazil (75)	India (130)

Selected diseases

We will consider two main group of diseases in the study, based on their prevalence in developed and developing countries. To this effect, we rely on the *Types* of diseases defined by the Consultative Expert Working Group on Research and Development: Financing and Coordination (CEWG) of the World Health Organisation. Three types of diseases are defined³:

- Type I: incident in both rich and poor countries, with large numbers of vulnerable populations in each
- Type II: incident in both rich and poor countries, but with a substantial proportion of cases in poor countries
- Type III: are those that are overwhelmingly or exclusively incident in developing countries

² <http://hdr.undp.org/en/2015-report>

³ WHO Secretariat. Defining disease types I, II, and III. (http://www.who.int/phi/3-background_cewg_agenda_item5_disease_types_final.pdf)

For our study, we will focus on a selection of diseases of Type I and Type III, given that the former affect more or less equally developed and developing countries, and the later affect specially developing countries. Type III diseases basically correspond to the NTD.

Table 2. Selected diseases by type

Type I	Type III
Ischaemic heart disease	Chagas disease
Diabetes mellitus	Leishmaniasis
Cerebrovascular disease	Schistosomiasis
Liver cancer	Onchocerciasis
Alzheimer	Trypanosomiasis

Identifying research on diseases

The identification of scientific research on specific diseases builds upon a previous study conducted at CWTS aimed at quantifying research outputs by disease. This was done coupling of publications to International Classification of Disease (ICD)-10 Diagnose Groups via keywords (manuscript in progress).

The Web of Science (WoS) was used to identify scientific research on different diseases, focussing only on biomedical research fields.

There were selected 84 out of the 250 WoS categories that are most medically oriented. The selection was validated looking at the research output of the eight Dutch university medical centers. The outcome of this validation exercise indicated that over 95% the publications was in one of the selected web of science categories. The dataset originally built included all articles and reviews in the 84 WoS categories, published between 2000 and June 2014. This original dataset contained 6.5 million publications in total.

In our study, we will select publications in the period 2009-2014, covering the most recent period. In this period we will be also able to analyse the origin of the funding used in research on specific diseases as acknowledged by authors in their publications

Scientific impact of journals and publications

The level of citation impact of journals and publications will be determined using CWTS' standard indicators. In the case of journals, we will divide each WoS subject category in four quartiles based on the Mean Normalized Journal Score (MNJS), which will allow us to distinguish between journals with a higher impact (1st and 2nd quartiles) and journals with lower impact (3rd and 4th quartiles).

Citation impact of publications will be determined mainly using the Mean Normalized Citation Score (MNCS).

PRELIMINARY FINDINGS

Table 3 shows the total number of publications produced by each country in the period 2009-2014 and the amount of publications dealing with specific diseases. In this table can be observed how very high developed countries devote a higher amount of publications to specific diseases, also Brazil, reaching comparable figures.

Table 3. Publications by country in the period 2009-2014

Country	All publications	Publications on diseases (%)
Netherlands	210,813	36,271 (17.2)
Spain	310,285	35,204 (11.3)

Russia	174,596	5,923 (3.4)
Brazil	220,803	31,583 (14.3)
Colombia	18,781	1,847 (9.8)
India	291,877	23,301 (8)

In table 4 we present an example of the amount of publications each country devoted to a disease of Type I (Diabetes) and a disease of Type III (Chagas disease). This table reveals that very high developed countries devoted much more research to Diabetes while the relative effort of countries like Colombia is much more concentrated on Chagas disease. Brazil, a country suffering both diseases and with a relatively good position to conduct research, presents a quite balanced profile.

Table 4. Publications devoted to Chagas disease and Diabetes

Country	Diabetes (%)	Chagas disease (%)
Netherlands	517 (1.4)	8 (0)
Spain	542 (1.5)	114(0.3)
Russia	125 (2.1)	5 (0.1)
Brazil	774 (2.5)	664 (2.1)
Colombia	17 (0.9)	82 (4.4)
India	689 (3)	4 (0)

REFERENCES

- European Commission (2013) Investing in health. Commission staff working document. Social investment package.
- Evans JA, Shim J-M, Ioannidis JPA (2014) Attention to Local Health Burden and the Global Disparity of Health Research. PLoS ONE 9(4): e90147. doi:10.1371/journal.pone.0090147
- Gillum LA, Gouveia C, Dorsey ER, Pletcher M, Mathers CD, McCulloch CE, et al. (2011) NIH Disease Funding Levels and Burden of Disease. PLoS ONE 6(2): e16837. doi:10.1371/journal.pone.0016837
- King, DA (2004) The scientific impact of nations. Nature 430, 311-316
- Kingel, JM; Roxrud, I; Vollset, SE; Skirbekk, V; Røttingen, JA (2014) Are the Norwegian health research investments in line with the disease burden? Health Research Policy and Systems, 12:64.
- Wilder, R; Solovy, EM (2005) The development of medicines for developing countries diseases: the role of intellectual property. International Seminar on Intellectual Property and Development, 2-3 May, Geneva (Switzerland)



Biodiversity Sustainability of Phytomedicine Research: a 3-dimensions analysis around the North-South divide.

Philippe Gorry*

*philippe.gorry@u-bordeaux.fr

GREThA UMR CNRS 5113, University of Bordeaux, Avenue Duguit, Pessac, 330608 (France)

INTRODUCTION

The use of plants in medicine can be traced to the beginnings of civilization and natural products dominated therapeutics until the end of the 19th century. The Industrial Revolution and the development of organic chemistry resulted in a preference for synthetic products for pharmacological treatments (Rates, 2001). However, according to the World Health Organization (WHO, 2008), about 65–80% of the population in developing countries depend on plants for their primary health care.

During the last decades, popularity of alternative medicines increased worldwide, especially phytomedicine. The global trade of medicinal plants was around \$62 billion in 2000's and expected to be \$5 trillion by 2050 (Kumari et al., 2011). The rising demand of plant-based drugs is creating heavy pressure on some plant populations in the wild due to over-harvesting, raising conservation and equity issues in regard to biodiversity and traditional knowledge (TK)(Arihan et al., 2007). Reducing the pressure on medicinal plants is therefore a tough challenge both for policy makers and economists (Timmermans, 2003).

The “Intellectual Property Rights” (IPR) linked to the use of plants have been debated worldwide and significant divergences exist as to whether IPR should be applied (IBC Working Group, 2010). To protect TK, there are two approaches: a positive protection route and a defensive approach route, with IP or non-IP related tools, legally binding or non-binding instruments. The debate on the right tools is not over (Van Overwalle, 2005). There are arguments for the benefit sharing under the IPR, considered as a new legal form of bio-piracy (Patil, 2012), whereas others argue that the IPR is a legal tool to protect the rights of knowledge holders and sustain innovation for the benefit of public health.

Since the adoption of the Convention on Biological Diversity (CBD), protection of the rights of local people and knowledge, as well as biological resources conservation are enforced (WIPO, 2015). Every member of the World Trade Organization has to comply with its agreement on the protection of IPR, and Trade-related aspects of IPR (TRIPS) agreement (Nair, 2011). Unfortunately, there is a wide gap between developed and developing nations on patenting natural products.

This debate is reminiscent of the social dilemma pinpointed by Hardin (1968) in his “Tragedy of commons” in relation to environment sustainability and protection, and the mirror-image of the concept, “the tragedy of the anti-commons” developed later on by Heller and Eisenberg (1998) to underline coordination failure induced by patent thickets. The concept was revisited by the Nobel Prize-winning economist, E. Ostrom (1999), because it might not be as prevalent. It could be desirable for a country to have an equal emphasis on conservation and market appropriation of medicinal plants. This can be facilitated through public-private partnership involving indigenous communities (Suneetha & Chandrakanth,

2006). In the past years, the World bank and the OECD have launch inclusive innovation actions, and countries such as India, China or South Africa have added inclusive element in their policies (Foster & Heeks, 2013).

Objectives

While the debate is still going on, no one's have assessed in the scientific literature or elsewhere, the output of the research and the extend of patent protection in phytomedicine, to get an understanding of the market organization, and relationships between firms and academia across North-South divide (Frickmann, 2011).

In order to contribute to the debate, we undertook a 3-dimensions analysis with various Science and Technology Indicators (STI) to analyze the biodiversity distribution, the research trends on medicinal plants, the patent landscaping and the competitive positioning of economic agents.

METHODS

Literature search on phytomedicine research was conducted in Scopus[®] database with keywords such as “Phytomedicine”, “Phytotherapy”, and “Medicinal Plants” in title, abstract or keyword fields. A collection of 20028 documents was retrieved. For the bibliometric analysis, authors, subject, affiliation, and country were deeply examined.

Patents were retrieved from the INPADOC family patents worldwide collection using Questel[®] software with the help of the International Patent Classification codes. We used the code A61K-036/00 and the subsequent sub-groups defined as “Medicinal preparations of undetermined constitution containing material from algae, lichens, fungi or plants, or derivatives thereof, e.g. traditional herbal medicines“. We identified 13489 patent families between year 1991 and 2014. Patent bibliographic data were analyzed with the Intellixir[®] software for the different information fields (priority or publication date and country, applicant name).

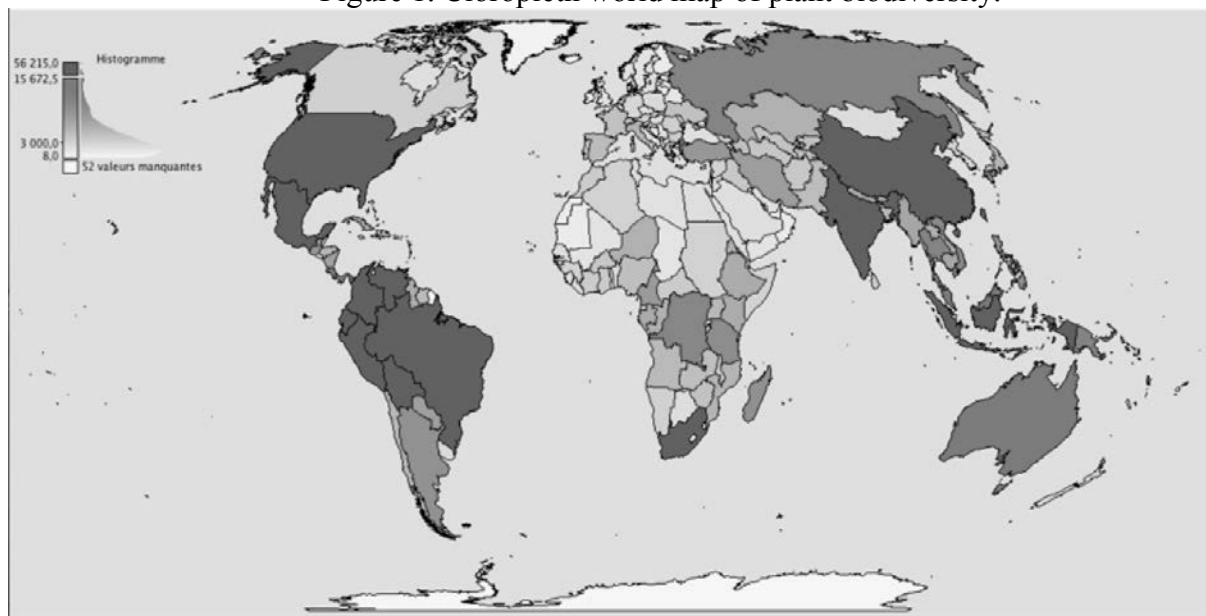
Descriptive statistics were run for the different indicators and some inferential statistics were performed using regression analysis or principal component analysis (PCA) with the help of XLSTAT[®] software in order to describe association or correlation between science, innovation and macroeconomic indicators. Science and technology indicators (STI) were gathered from the OECD STI statistics web portal. All analysis, metrics and statistics were performed on patent family. Choropleth map was draw to visualize some statistical variable on geographical area with the help of MAPresso[®] web application (Herzog, 2003).

RESULTS

Geography of Biodiversity

To address the issue to which extend the biodiversity is pre-empted by IPR, we look in the scientific literature for data on plant taxon total number of know species as well estimate of unknown species and their geographical distribution (Joppa, Roberts & Pimm, 2010). It is remarkable to observe that biodiversity is share between OECD and BRICS countries: Australia, Brazil, China, India, Indonesia, Malaysia, South Africa, USA are the major biodiversity hotspots in the world (Fig. 1).

Figure 1. Choropleth world map of plant biodiversity.

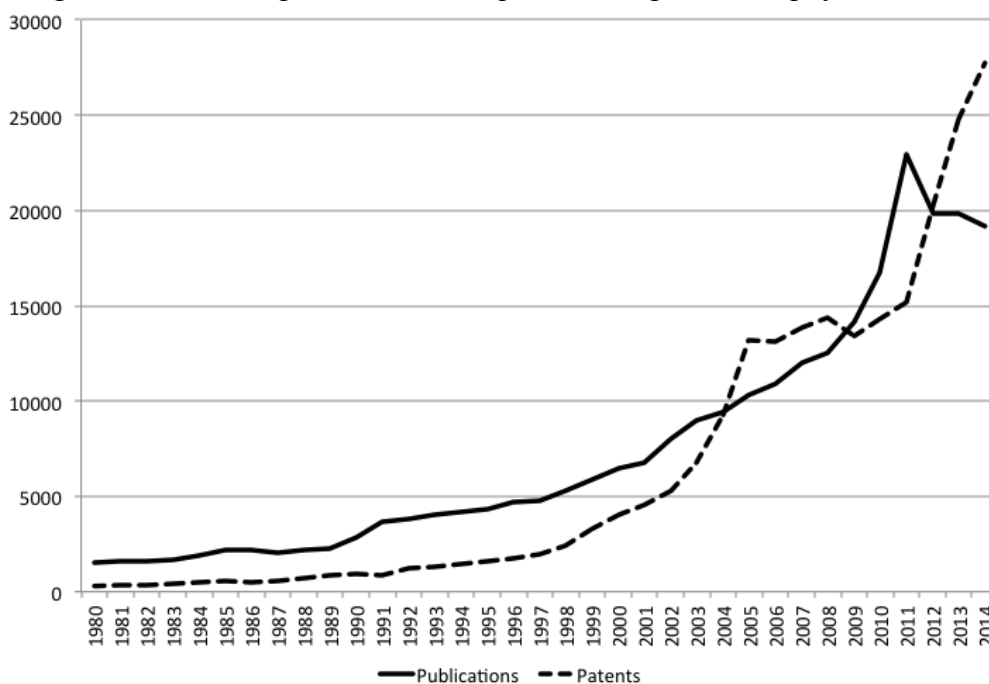


However, if we look at the geographical distribution of estimated unknown species, we got a different picture with Australia, Central and South America, and South Africa the main reserve areas (data not shown).

Research in phytomedicine

Scientific publications trend was calculated and compare to patent filling trend between 1990 and 2010 in the field of phytomedicine (Fig.2).

Figure 2. Scientific publications and patent filling trends in phytomedicine.



Despite to be considered as an old scientific field, it's an active research domain producing more than 3000 publications by year; scientific publications exceeded patent applications by year 2003.

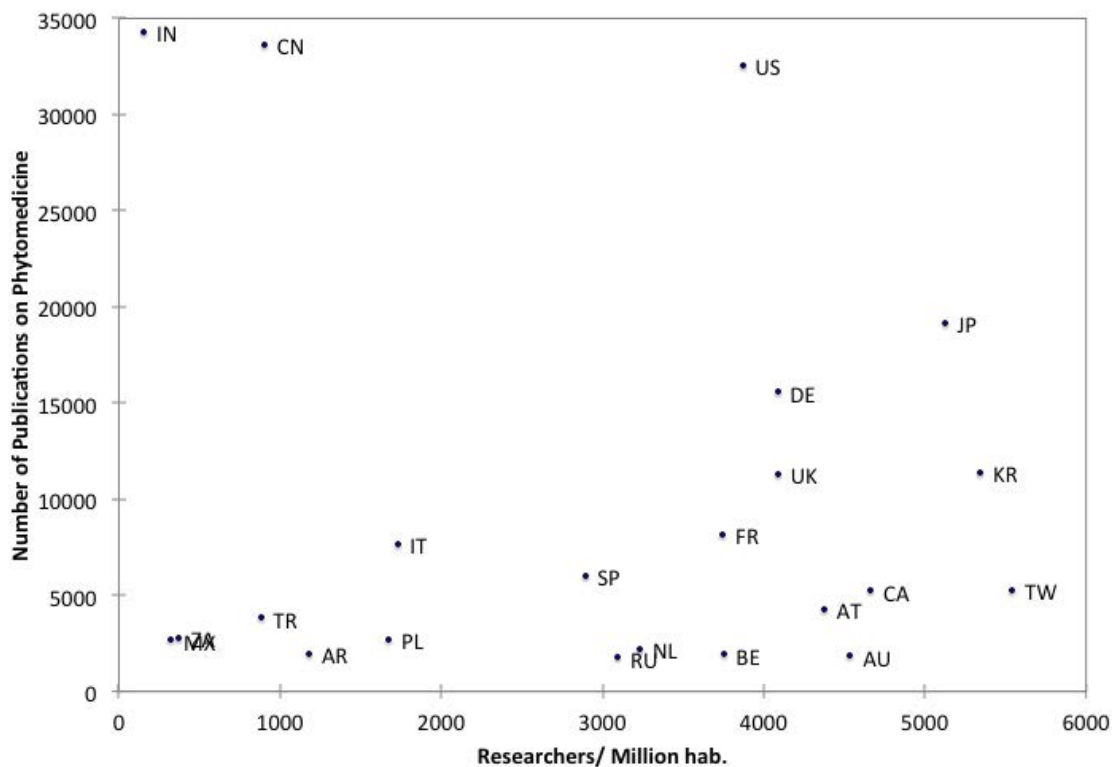
Then a ranking of the academic institutions publications was draw based on publications number (Tab. 1).

Table 1. Ranking of top academic institutions by country.

Ranking	Affiliation	Publications	Country
1	Chinese Academy of Medical Sciences	1911	CN
2	Universidade de Sao Paulo - USP	1800	BR
4	Kyung Hee University	1533	KR
7	University of Karachi	1312	PK
10	University of Toyama	1105	JP
15	Technische Universitat Berlin	930	DE
18	National Research Centre	892	EG
19	University of Illinois at Chicago	889	US
21	Universiti Putra Malaysia	865	MY
24	Universite de Yaounde I	811	CM

Interesting, many of the top research centers in phytomedicine are located in BRICS countries, and they are highly competitive, especially in comparison with US universities.

Figure 3. Phytomedicine research productivity by country.



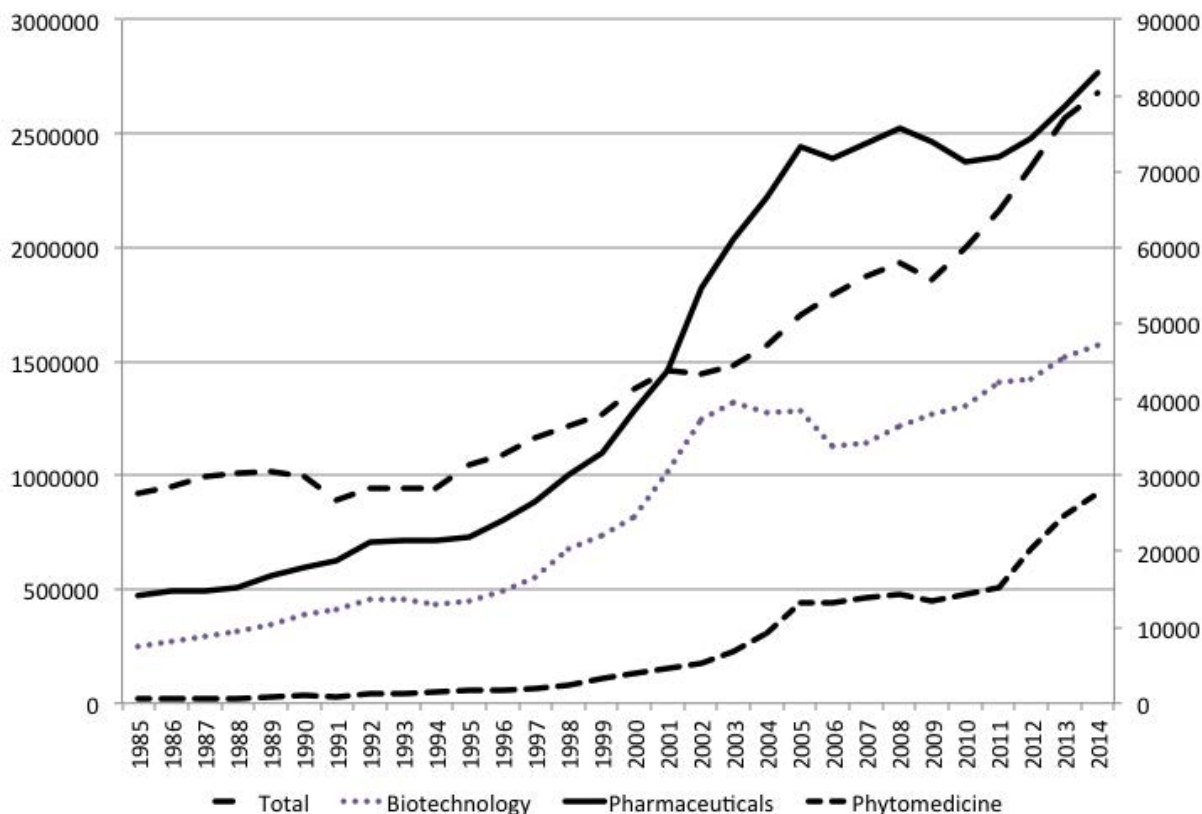
Among the top 100 research institutions, OECD account for 43% (US weighting 8% and EU 14%), BRICS 36% and other developing countries 21%.

We look at the productivity of this scientific field in relation with the size of the population country (data not shown) or the number of PhD and engineers in a given country (Fig. 3). BRICS countries such as Brazil, China, or India show a very competitive ratio compare to most OECD countries including USA.

Patenting medicinal plant extract

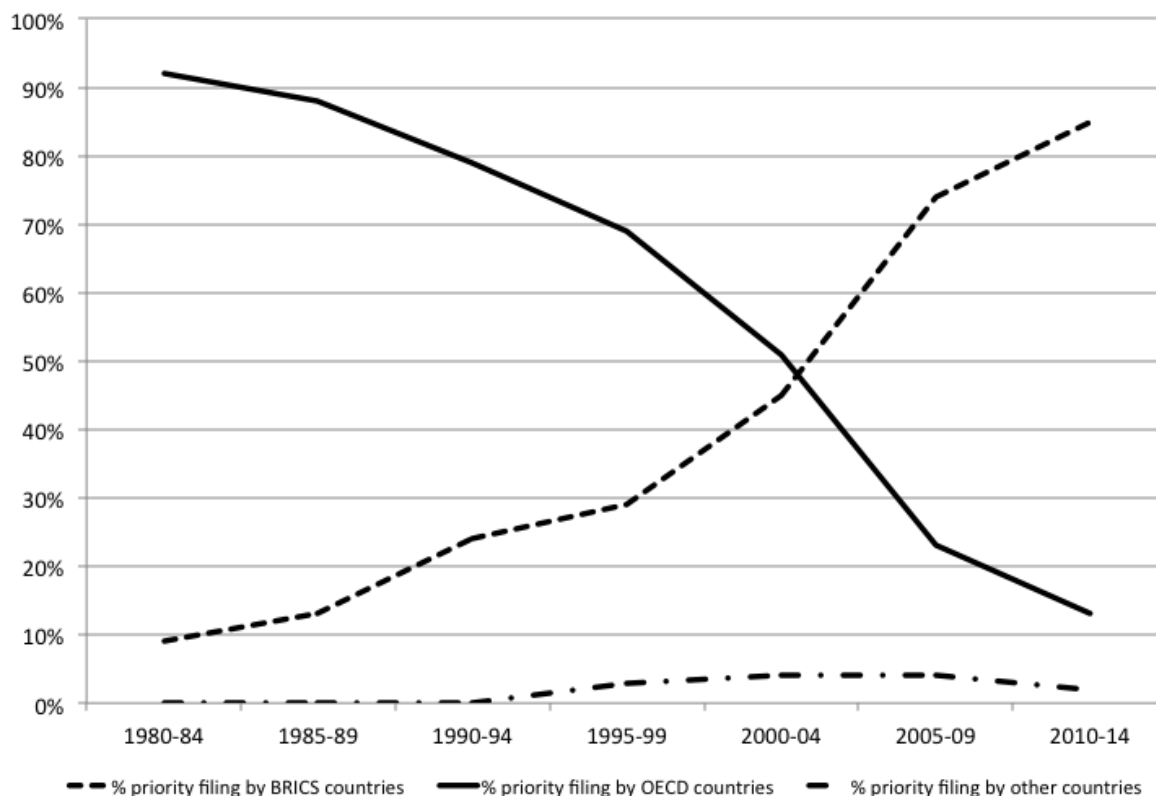
Patent filing trend shows that phytomedicine has undergone a tremendous increase in IP protection followed by an accelerated worldwide IPR strategy. 72,5% of patents has been file since 2000 with more than 1000 new applications every year. It is following biotech and pharmaceutical compound trends. But this steady increase stops during the years 2005-2010: it coincides with the IPR negotiation round and the issuance of UN declaration on the rights of indigenous people in 2007 (Fig. 4).

Figure 4. Patent filling trends for biotechnology, pharmaceutical & phytomedicine inventions compared to all patent filling.



Phytomedicine patenting is highly concentrated with 68,80% of the patent family filed by China. If OECD countries accounted for 70% of the total patents between 1995 and 2000, BRICS have caught up around 2005, and represent today more than 80% of total patent filing while OECD new patent filing decrease by a third (Fig. 6).

Figure 6. Phytochemistry patents filing by country or group of country.



Looking at the scientific production, China, India and Brazil are very effective in regard of their number of academic researchers or percentage of GDP dedicated to R&D. But, only China seems to have an IPR strategy with inclusive innovation elements. According to the balance measure between priority filing (or PCT, Patent Cooperation Treaty) and patent publication, we can observe that China move from a position where it was the target of foreign assignees filing PCT to a position where Chinese firms and academia are protecting their internal market and extend increasingly their IPR worldwide through PCT (Fig. 7).

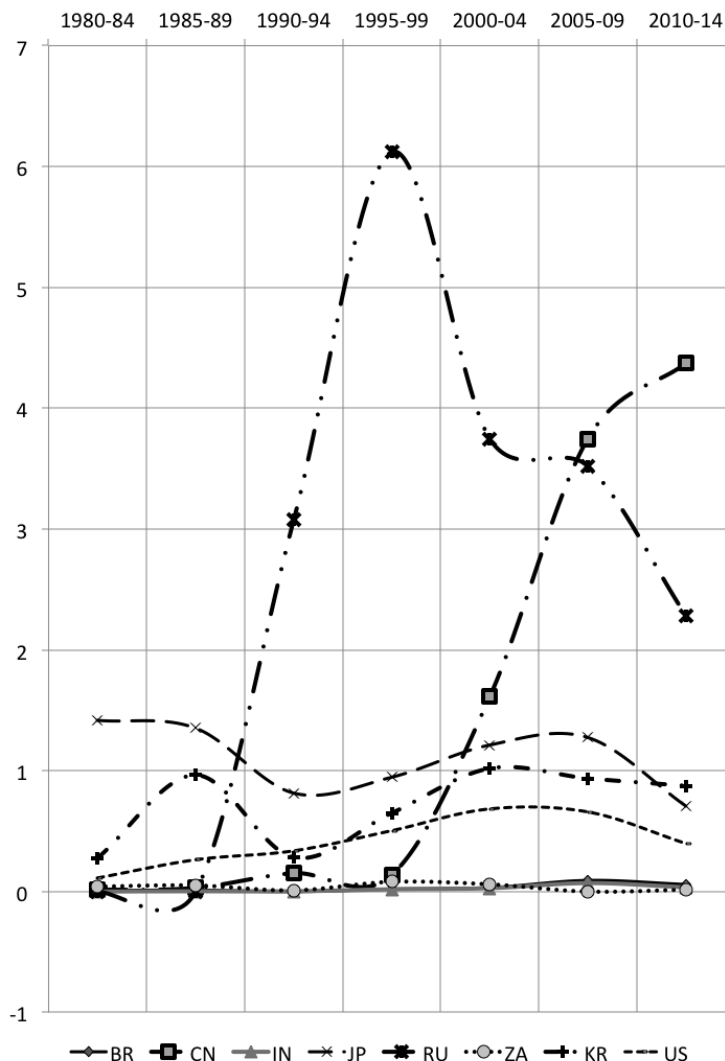
During the last decades, most of the patents claim therapeutic effects against skin diseases with new trends on osteoporosis, diabetes & anti-oxidants. The innovation trend on anti-oxidants is so important that there is more patent filed today on plant extracts than on pharmaceutical compounds for this therapeutic class (data not show).

As you will expect BRICS countries are more focus on innovation addressing their public health needs but it's not the case for all developing countries. Therapeutic classes with the highest international extension of IPR (PCT applications) are plant extracts with immunomodulation property, especially anti-psoriasis. The main therapeutic area patented in BRICS countries is eupeptics for the digestion (Viegas, 2007; data not shown).

Whatever the medical applications, most of the patent claims IPRs on extracts from dycotyledons with Magnoliopsida, one of the largest plant taxa. For the past decade, new patent filings were focusing on extracts from Rosaceae, or more recently on Curcuma. The plant specie targeted by the biggest number patent filing number is the Theaceae (tea family). OECD and BRICS patent applicants have different biodiversity interests. The plant species subject to the strongest international IPR protection are Vitaceae or Ampeliidaceae (example: vitis rotundifolia), with Gynkgophyta, the plant specie, targeted by triadic patent, while

Glycyrrhiza specie is mostly protected by BRICS (Frickmann, 2011; data not shown).

Figure 7. Evolution between 1980 and 2010 of the balance of IP share in BRICS countries.



Looking at the patent assignees distribution, it is not a surprise to find that 90% of them are private companies. Historically, it was mainly Japanese firms, with the exception of one academic institution the Indian Council for Scientific and Industrial Research (CSIR).

At the present time (years 2010-2014), the picture is completely different: Chinese applicants are dominating the scene with 63% of the top 100 patent applicants, followed by Korean (12%) and Japanese (3%). While in the past academic institutions were marginal in filing patents, they represent 46% of the top 100 applicants during the years 2010-2014.

Among the top applicants, are present European global companies and Asian firms, in cosmetic or health nutrition, and academic institutions, mainly from China. To be noticed, Indian or South African, firms or academia, are absent of the top ranked applicants. There is no monopoly on IPR since the top 10 assignees represent around 1% of the patent filing (Tab. 2).

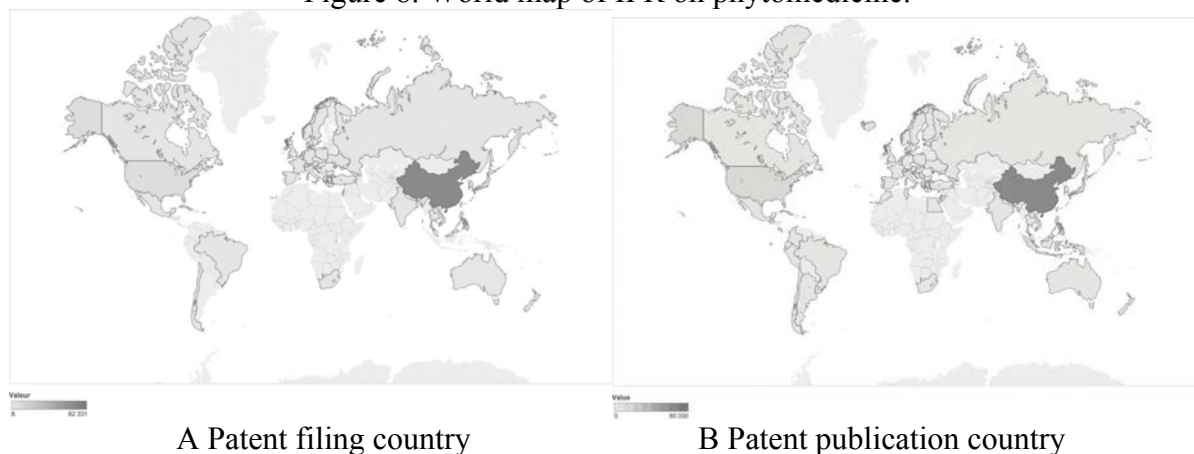
Table 2. Ranking of applicants by patent filing between 2010-2014.

Rank	Applicants	Patent filing	Status	Country
1	Amorepacific	253	Private	KR
2	Qingdao ShizhiliInstitute of Chinese Herbal Medicine	186	Public	CN
3	Jinan Weichuan InfTechnology	178	Private	CN
12	Kao	146	Private	JP
16	Kyung Hee University	115	Public	KR
19	National University of Pharmacy	112	Public	UA
26	Shiseido	92	Private	JP
30	Nestec	85	Private	CH
55	FANCL	52	Private	HK
67	DSM	46	Private	NL
81	US National Institutes of Health	41	Public	US
90	Federal Noe G Bjudzhetnnoe Obrazovatel ...	39	Public	RU
96	L'Oréal	36	Private	FR

On the overall, phytomedicine innovations are undergoing a worldwide strategy of IP protection by numerous country and firms as a comparative choropleth world map of priority and publication filing can show in Figure 8. Patent applications around published all around the world with the notable exception of sub-Saharan Africa.

Despite the fact that China is one of the top filing country, Chinese applicants don't have a strong international IPR strategy: less than 10% of their patents undergo a PCT filing (data not show).

Figure 8. World map of IPR on phytomedicine.



Finally, we analyze the patent co-applicant network in order to look at the collaborative pattern of this market. It is characterized by a small-world pattern of collaboration without any concentration and monopoly build around global company. Few collaborative networks between firms and academic institutions are observed like between the French public research organism CNRS and Nestle, Tianjin University and Tasly Pharmaceuticals or the US NIH and Shiseido Japanese international company (data not show).

DISCUSSION

These results highlight the debate around IPR and medicinal plant extracts. Research on phytomedicine has undergone a tremendous increase followed by an accelerated worldwide patent protection strategy. Three quarters of patents have been filed since 2000 and the different rounds of IPRs negotiation around biodiversity (CBD and TRIPS). This patenting trend is following the one observed for biotechnology. It is sustained by an active academic research led by BRICS countries. China and India are very effective in regard of their number of academic researchers or % of GDP dedicated to R&D. Globally, BRICS and developing countries are leading today phytomedicine research. But, only China seems to successfully translate research discoveries in protected inventions.

The IPR pre-emption might be linked to the presence of global companies. The European or Japanese firms are not responsible of patent thicket in this market: they don't have any worldwide IP protection monopoly. In fact, the pre-emption of biodiversity through IPR by firms from OECD countries is challenged by BRICS scientific research output.

OECD countries past monopoly is moving away with new player like China which succeed in implementing successful policies, linking science and technology transfer (Jiang, 2011), while India is still behind despite public policy in favor of TK and reverse medicine national research program such as the New Millennium Indian Technology Leadership Initiative scheme (Patwardhan & Mashelkar, 2009). The danger of economic monopoly regarding TK is still remaining especially for dermato-cosmetology medicinal claims, but not for the overall medical applications since it seems that big pharma have give up this research field.

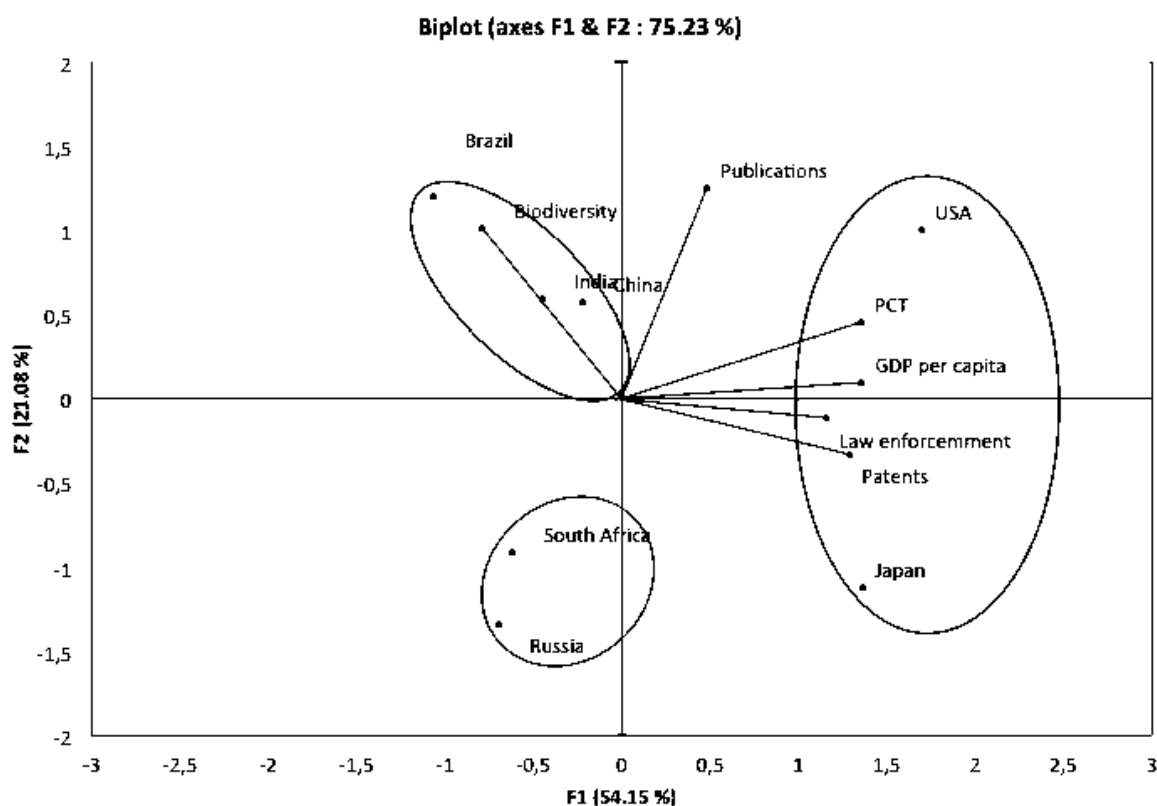
It remains to understand, which of these strategies, the positive protection through patent like in China or the defensive approach by disseminating indigenous knowledge, like in India is more efficient against acquisition and exploitation by third parties (Van Overwalle, 2005).

There are many opportunities for public-private partnerships in a highly competitive and innovative sector with numerous new firms. Innovation public policies should encourage academic institutions to protect their inventions in order to prevent private monopoly and foster translational research, and North-South translational collaborative research should be sustained (Dinopoulos & Segerstrom, 2010). Importance of interaction between university, industry and government in innovation system is stressed by the concept of Triple Helix model of innovation system (Etzkowitz and Leydesdorff, 1995). Public R&D in developing countries as to fill the gap but appropriation of R&D gains is difficult because IPR enforcement is a challenge in developing countries.

In order to explore endogenous and exogenous factors, we undertook a PCA as multivariate analysis technique to find which variable are more appropriate to describe association between biodiversity, phytomedicine research (Publications), IPR pre-emption (Patents and PCT), macroeconomic indicators (GDP per capita) or regulations (law enforcement) in the context of national public policies (Vyas, 2009). Multiple linear regression models will be necessary in future research to establish any correlation (Fig. 9).

In conclusion, this patent mapping will improve the overall transparency of phytomedicine IPR for researchers, policy-makers and the civil society. Furthermore, these results will provide technology transfer strategic information for the academic research and foster successful open innovation partnership model.

Figure 9. Principal Component Analysis biplot (arrows: variables; dots: observations).



REFERENCES

- Arihan, O. & Gençler Özkan M. (2007), Traditional medicine and Intellectual Property Rights. *J. Fac. Pharm. Ankara*, 36: 135 -151.
- Dinopoulos, E. & Segerstrom, P. (2010). Intellectual property rights multinational firms and economic growth. *Journal of Development Economics*, 92, 13-27.
- Etzkowitz, H. & Leydesdorff, L. (2002), Universities and the Global Knowledge Economy: A Triple Helix of University-Industry-Government Relations, New York: Continuum International Publishing Group.
- Hardin, G. (1968), "The tragedy of the commons". *Science*, 162, 1243-1248.
- Frickmann, S. (2011). Research & Patent of Phytotherapeutic & Phytocosmetic Products in the Brazilian Amazon. *J. Technol. Manag. Innov.*, 6: 136-150.
- Foster, C. & (2013), Conceptualising Inclusive Innovation: Modifying Systems of Innovation Frameworks to Understand Diffusion of New Technology to Low-Income Consumers. *Eur. J. of Development Research*, 25, 333-355.
- Heller, M. & Eisenberg, R.S. (1998). Can patents deter innovation? The anticommons in biomedical research. *Science*, 280, 698-701.

Herzog, A. (2003). Developing Cartographic Applets for the Internet. In MP. Peterson (Ed.), *Maps and the Internet, 1*: 115-128.

IBC Working Group. (2010). Draft preliminary report on traditional medicine & its ethical implications. Paris: UNESCO.

Jiang, J. (2011). Protecting China's national treasure. *Nature*, 480: S93.

Joppa, L.N., Roberts, D.L., & Pimm S.L. (2010). How many species of flowering plants are there? *Proceedings of the Royal Society, B: Biological Sciences*, 278: 554-559.

Kumari, S., Shukla G. & Rao A.S. (2011) The Present Status of Medicinal Plants Aspects and Prospects. *International Journal of Research in Pharmaceutical and Biomedical Sciences*, 2: 19-22.

Nair, M.D. (2011). TRIPS, WTO and IPR: protection of bioresources and traditional knowledge. *Journal of Intellectual Property Rights*, 16, 35-37.

Ostrom, E. *et al.* (1999). Revisiting the Commons: Local Lessons, Global Challenges. *Science*, 284, 278-282.

Patil, D.A. (2012). Indian Ethnobotany: genesis through ages. *World Journal of Science and Technology*, 2, 01-14.

Patwardhan, B. & Mashelkar, R.A. (2009). Traditional medicine-inspired approaches to drug discovery: can Ayurveda show the way forward? *Drug Discovery Today*, 14, 804-811.

Rates, S.M.K. (2001). Plants as source of drugs. *Toxicon*, 39: 603–613 .

Suneetha, M.S. & Chandrakanth, M.G. (2006). Establishing a multi-stakeholder value index in medicinal plants—an economic study on selected plants in Kerala and Tamilnadu States of India. *Ecological Economics*, 60, 36- 48.

Timmermans, K. (2003). Intellectual property rights and traditional medicine: policy dilemmas at the interface. *Social Science & Medicine*, 57, 745-756

Van Overwalle, G. (2005). Protecting and sharing biodiversity and traditional knowledge: Holder and user tools. *Ecological economics*, 53, 585-607.

Vyas S. & Kumaranayake L., (2006), Constructing socio-economic status indices: how to use principal components analysis, *Health Policy and Planning*, 21, 459-468.

World Intellectual Property Organization, Intellectual Property and traditional knowledge, Publication No. 920(E), Geneva: WIPO.

World Health Organization (2008), Fact sheet: traditional medicine., Geneva: WHO from: <http://www.who.int/mediacentre/factsheets/2003/fs134/en/>



***In Re*ⁱ The Academic Cartography of Sugar Sweetened Beverages: Scientific and Technical Information, Interdisciplinarity, and Legal Academia**

Lexi C. White, J.D., Ph.D. Candidate *

*acwhite5@asu.edu

School of Human Evolution and Social Change, Arizona State University, 900 Cady Mall, Tempe, Arizona 85281 (United States)

INTRODUCTION

It has long been noted that there are crucial differences between the functions of the law and science (Brief for Bloembergen et al., 1993). Difference in function leads to different logics and processes of enacting law and science, including the system for generating a body of peer-reviewed research. The current research uses an overarching anthropological approach to address academic communication between two powerful and influential social groups: scientists and legal scholars. Using Burt's (1992) structural holes, which examines the position of actors across network gaps, and the newer area of cultural holes, which adds a cultural dimension through linguistic networks (Pachucki & Breiger, 2010), this research looks at whether, in addition to structural divides in patterns of citations between legal academic and scientific publishing, there are also language and contextual differences. By investigating these key issues this research will not only expand the applications of these well validated scientometric techniques to new areas, but also will explore the intersection of two academic publication areas, the way they communicate, and how information across both is attempting to influence policy.

THEORETICAL BACKGROUND

One of scientific research's most complex problems is generating on-the-ground, real-world impacts based on scientific and technical research results (National Research Council, 2012). These issues have come to the forefront recently, particularly in the United States, given the uses and misuses of scientific and technical information outside academic walls in the hands of climate-change deniers, anti-vaccination movements, and even bans on soda. Fears about poorly executed or even fabricated research being used to direct not only public hysteria, but also actual policy decisions are not wholly unfounded.ⁱⁱ Politicians and regulatory agencies are swayed not only by journalists and public outcry, but also by a very particular subset of educated academics who structure their research specifically to influence statutes,ⁱⁱⁱ regulations, and court decisions^{iv}: legal academic scholars.

Legal scholars operate in a publication domain parallel to, but distinct from, scientific scholars. Their articles go through a similar (but student driven) peer-review process, and are published in journals accessible through different databases (also owned by Thompson Reuters and Elsevier). Legal scholars also extensively pull evidence from across scientific fields, potentially acting as an information bridge between scientists and policy makers.

SUGAR SWEETENED BEVERAGE PUBLICATIONS

While existing research has established that structural holes of differing depths exist between most scientific disciplines (Leydesdorff & Rafols, 2009), this research includes legal academic journals, which were not included. This research predicts that while there will be gaps between clusters of papers from legal journals and papers from scientific journals, there will be a large number of legal papers that occupy a bridging position between other legal papers and scientific papers because legal authors frequently cite to scientific papers but not vice versa. Even though structurally, legal papers will be linked to scientific papers, the linguistic content of legal papers will be significantly different and distant from that of scientific papers, because legal authors are more likely to use prescriptive rather than descriptive language.

The current dataset for this research focuses on one key public health policy area (Studdert et al., 2015) and contains all articles retrieved using the search terms “sugar sweetened beverages” on both SCOPUS and LexisNexis^v. At the time of this submission, the article data from both of these databases is being cleaned and combined into a format that allows for analysis.

Data Analysis

After data cleaning and processing is complete, the researcher will create a cocitation matrix that shows when papers in the network have cited one another. Multidimensional scaling will then be used to generate a two dimensional representation that places like papers nearer to one another in order to visualize the network of sugar sweetened beverage papers. The underlying cocitation matrix of papers will then be used to calculate the structural holes, or gaps between clusters of similar papers or authors. In the context of a cocitation network, those occupying bridging positions between clusters of similar authors are those who cite across disciplines and engage in what may be interdisciplinary research (Porter et al., 2007). By calculating the structural holes between clusters, and, thus, the positions of the brokers in the cocitation network, investigators can identify these structurally interdisciplinary papers and authors. Given that legal authors frequently cite scientific papers, but the reverse is not true, it is likely that a structural network will show that legal papers frequently occupy broker positions across structural holes between scientific and legal journals.

Recently, networks of structural holes have been combined with linguistic networks in order to investigate cultural clusters and divisions as well (Pachucki & Breiger, 2010). In addition to looking at the citation patterns between articles, textual content will be probed using latent Dirichlet allocation (LDA) in order to generate a distribution of topics, or salient word categories, for each article. LDA allows for the text itself to generate a series of topics that may indicate not only different types of prescriptive or descriptive language as more associated with legal or scientific papers, but also other emergent themes that may be strongly associated with one area. For example, legal scholars writing on sugar sweetened beverages may frequently compare SSBs to tobacco in a way that scientists do not.

Once topic distributions are generated for each article, a matrix of the distribution of topics across documents will be used to calculate the gaps between clusters of articles providing a cultural hole analysis. This analysis will likely show that while structurally many legal papers occupy a broker position between legal and scientific frameworks, their textual context is distinct and clusters exclusively with other legal academic papers.

SIGNIFICANCE OF RESEARCH

By examining how legal academic publications fit into networks of scientific publications in terms of paper and author citations this research will expand the science of science to cover new disciplines and new databases. This research will blend a variety of research frameworks from anthropological to sociological, to legal and a variety of research methods in order to investigate a rarely studied transdisciplinary interaction. This research also will go beyond cursory exploration of most-cited authors, additionally exploring network patterns with multidimensional scaling and structural holes, as well as looking at content through cultural hole analysis.

The proposed research will not only provide key insight about legal education and legal academic processes, but it will also advance the fields of scientometrics by applying existing cocitation network techniques to novel academic areas and novel databases. While scientometrics frequently looks at documents generated from legal processes, such as patent applications, this research expands the purview of citation analysis to legal academic databases, working on not only new techniques for gathering and processing data, but also expanding the existing maps of academic literature.

REFERENCES

- Belware, K. (June 4, 2015). Doctors Sue Arizona Over Abortion Law that ‘Forces Doctors to Lie to Patients’. *The Huffington Post*. Retrieved from: http://www.huffingtonpost.com/2015/06/04/arizona-abortion-lawsuit-_n_7513688.html.
- Brief for Bloembergen, N., et al. as Amici Curiae, *Daubert v. Merrell Dow Pharmaceuticals Inc.*, (1993) 509 U.S. 579. Retrieved from: <http://www.atlanticlegal.org/daubert.html>.
- Burt, R. S. (2015). Reinforced Structural Holes. *Social Networks*, 43, 149–161.
- Burt, R. S. (1992). *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.
- Closen, M. L., & Dzielak, R. J. (1996). History and Influence of the Law Review Institution, *Akron Law Review*, 30, 15.
- Delgado, G. & Davenport, M.L. (2012). Progesterone Use to Reverse the Effects of Mifepristone. *Annals of Pharmacotherapy*. 46(12).
- Leydesdorff, L. & Rafols, I. (2009) A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology*, 60, 348-362.
- National Research Council. (2012). *Using Science as Evidence in Public Policy*. Washington, D.C., The National Academies Press.
- Newton, B. E. (2012). Law Review Scholarship in the Eyes of the Twenty-First Century Supreme Court Justices: An Empirical Analysis. *Drexel Law Review*, 4, 399.
- Pachucki, M. a., & Breiger, R. L. (2010). Cultural Holes: Beyond Relationality in Social

Networks and Culture. *Annual Review of Sociology*, 36(1), 205–224.

Porter, A. L., Cohen, A. S., Roessner, J. D., & Perreault, M. (2007). Measuring researcher interdisciplinarity. *Scientometrics*, 72(1), 117-147.

Studdert, D. M., Flanders, J., & Mello, M. M. (2015). Searching for Public Health Law’s Sweet Spot: The Regulation of Sugar-Sweetened Beverages. *PLoS Medicine*, 12(7).

ⁱ “*In re*” means “in the matter of.” A legal designation used for proceedings where there are no adversarial parties. In philosophy, the phrase has additional connotations of being about reality and the real world, as opposed to an ideal one.

ⁱⁱ For example, a recent law passed in Arizona that requires doctors to inform patients seeking a medical abortion that the procedure is reversible, based on a single study with only 6 participants (Delgado & Davenport, 2012) that has been rejected by the American Congress of Obstetricians and Gynaecologists (Bellware, 2015).

ⁱⁱⁱ Statutes and regulations rarely, if ever, provide citations, and so the causality between legal academic publications and law-making is difficult to determine. However, legal academic articles frequently explicitly advocate for new statutes and regulations, or changes to existing ones, and there have been a few cases where legal academic articles were known to influence statutory or regulatory changes, such as the influence of a 1982 Chicago Kent Law Review article on changes to the Illinois criminal code (Closen & Dzielak, 1996).

^{iv} A 2012 article found that “nearly” 2000 published U.S. Supreme Court opinions between 2001 and 2011 cited at least one or more law review articles (Newton, 2011).

^v LexisNexis is a database for American government, policy, and legal academic documents that is designed to assist legal professionals in their research.



Metrics and Agricultural Science - measuring Multidisciplinary and Applied Research

This note proposes a special session on agricultural sciences at the [STI Conference](#) (21st International Conference on Science and Technology), Valencia, 14-16 September 2016.

If we focus on the agricultural field, we see a kaleidoscopic picture. Agriculture includes a wide variety of economic activities, ranging from crop husbandry to cattle breeding and industrial processing of non-food products. It is often used in a broad sense to include for example forestry, aquaculture and fisheries. Agricultural sciences use methods from a wide variety of disciplines ranging from sociology to genomics. Although agricultural sciences are applied sciences there is a gamut from more fundamental studies to understand underlying processes to applied work to produce results that can be used directly in agricultural practice.

This complex picture raises a number of questions where bibliometric and altmetric indicators are concerned:

- Do agricultural scientists show publication and citation behaviours similar to other agricultural scientists, or are these behaviours more comparable to their disciplinary peers? For example: do rural sociologists follow similar patterns as agricultural economists or as sociologists in general? Do soil biologists publish and cite in similar ways as animal ecologists, or as other soil scientists?
- How well do the baselines often used in bibliometric studies (usually derived from Web of Science or Scopus) represent the agricultural sciences in view of its complexity?
- There is an assumption on the more 'fundamental' side of the gamut journal articles are used predominantly to communicate scientific results, while on the more 'applied' side other types of publications (like reports) play an imported role. How can applied and fundamental groups be compared?

The session will present the ongoing work in the field of indicators in different important institutions in the agricultural sciences.

Institut National de Recherche Agronomique-Laboratoires (INRA) Use Case

For 10 years, INRA bibliometric team has produced an annual report and thematic studies based on bibliometric indicators used by the scientific management to enlighten scientific strategy or give facts for the assessment of Inra academic performance. Who are the leaders in a specific domain? What are the strengths and weaknesses of Inra? How to measure inter-disciplinarily, how to assess applied research, how to compare properly the impact in different scientific disciplines? Could new metrics be useful for our domain? These questions are some of the examples that justify the use of a wide gamut of indicators, size dependent or not, considering long time period for our studies. It also implies a deep understanding of the organization of the bibliometric databases and of the domain to analyse to avoid false interpretation. This talk will give concrete examples of the use of indicators, a critical view of the results and some perspectives in a context of Open Science.

About the panellist

Vanessa Méry is a documentalist at INRA since 2007. She is part of the Scientific and Technical Information team and specializes in bibliometrics at the Jouy-en-Josas center. She obtained a M.Sc. in Ecology in 1997 at the Université de Créteil, followed by an MA in document engineering in 2005 at the CNAM. Since 2010, she annually performs bibliometric indicators for the INRA headquarters. In 2015, she started to co-manage the INRA's bibliometrics group, within the Scientific and Technical Information unit. Finally, she maintains an INRA bibliometric internet tool, NORIA, that allows to evaluate the impact of the scientific literature.

Soizic Messiaen holds a Msc in systematics evolution with a specialization in informatics applied to biology. She has been working at INRA since 2009, where she is responsible of INRA data and its normalization in the Web of Science. Since 2011, she is the co-administrator of the bibliometric tool NORIA that helps the interpretation of the impact factor and annually performs bibliometric indicators for INRA headquarters. Since 2015, she also co-animates the INRA bibliometric study group.

About INRA

INRA www.inra.fr/en is Europe's top agricultural research institute and the world's number two centre for the agricultural sciences. Its scientists are working towards solutions for society's major challenges.

Wageningen UR Use Case

Since 2008 WageningenUR has its own portal showing bibliometric indicators for individual researchers, groups or institutes. The used method, based on Web of Science and the Essential Science Indicators, has disadvantages for fields like Animal Science and Social Sciences. Shifting to tools like Incites or SciVal is very attractive, because of all extra features. In the project "Evaluation Bibliometric Analysis" we compare the various systems with our own method. In this project we try to get an answer on questions like: "Does an increase in research field affect field normalization?" and "Does a shift to Scopus lead to a better citation score for certain groups?" In this talk we will discuss the results of this project.

About the panellists

Hugo Besemer has been active all his working life in agricultural information, in Wageningen and the Dutch system but also for international and development organisations. He is now in the research support unit of the library and is concentrating on research data management, and innovations in the scholarly information systems like identifiers and altmetrics.

Ellen Fest holds a PhD in Soil Chemistry from Wageningen University. After that she worked during seven years for a Dutch engineering firm in the field of soil pollution and data analysis. In 2014 she switched to the library of Wageningen UR to specialize in bibliometrics and research data management. In 2015 she led the project in which 80 chair groups were evaluated on their bibliometrics.

About Wageningen UR

Wageningen UR is a consortium of the university and a number of specialized research institutes like Plant Research International and Alterra (environmental sciences). It ends up as one of the top institutions in university rankings in subject areas like agricultural sciences. Bibliometric information is routinely used for management procedures like research assessment exercises and tenure track appraisals

FAO of the United Nations Use Case

CAB Abstracts is a bibliographic database with the coverage of over 8.4 million records from 1973 onwards in applied life sciences includes agriculture, environment, veterinary sciences, applied economics, food science and nutrition provided by CABI (Centre for Agriculture and Biosciences International). It provides an international content including less well-known, non-English journals in 50 languages from over 120 both industrialized and developing countries. FAO use case aims at comparing the CABI database with main bibliometric databases (Scopus and Web of Science) in order to investigate: if they present bias in favour of topics relevant to industrialized countries; and if they present a coverage bias in favour of journals from industrialized countries.

About the panellist

Imma Subirats Coll has been working as senior knowledge and information management officer at the Food and Agriculture Organization of the United Nations (FAO) since 2006. She works on the advise of standards, tools and good practices for the management and exchange of data to academic, research, private and governmental institutions worldwide. She is also actively promoting open access and open data in the agricultural research context. In recent years, she has been working on the facilitation of the AIMS community and portal, space for accessing and discussing information management standards, tools and methodologies with the objective to connect information specialists worldwide to support the implementation of structured and linked information and knowledge.

İlkay Holt holds an MA in library science as well as an Executive MBA. She has been in the library profession for 17 years and held the position of Director of Libraries at Özyeğin University, Istanbul, between 2011 and July 2016. She is the public lead for Creative Commons Turkey and since 2015 has been a member of TÜBİTAK (The Scientific and Technological Research Council of Turkey) Open Science Committee. Holt was a Steering Committee Member at ANKOS from 2011 to May 2016 and during the same period she was the coordinator of the Licensing Group and advisor to the OA Working Group. Since 2005, she has taken up a variety of roles in Open Access such as E-LIS (E-prints in Library and Information Science) editor for Turkey, founding member of the Open Access Working Group of ANKOS (Anatolian University Libraries Consortium in Turkey), as well as working in the OpenAIRE Turkish National Helpdesk team. Since August 2016, she has been a freelance consultant in information management for COAR (Confederation of Open Access Repositories).

About FAO of the United Nations

The Food and Agriculture Organization of the United Nations (FAO) is a specialised agency of the United Nations that leads international efforts to defeat hunger. Serving both developed and developing countries, FAO acts as a neutral forum where all nations meet as equals to negotiate agreements and debate policy. FAO is also a source of knowledge and information, and helps developing countries and countries in transition modernise and improve agriculture, forestry and fisheries practices, ensuring good nutrition and food security for all. Its Latin motto, *fiat panis*, translates into English as "let there be bread". As of 8 August 2008, FAO has 191 members states along with the European Union, Faroe Islands and Tokelau which are associate members. It is also a member of the United Nations Development Group.



Network analysis to support research management: evidence from the Fiocruz Observatory in Science, Technology and Innovation in Health

Bruna de Paula Fonseca e Fonseca^{*}, Ricardo Barros Sampaio^{**}, Marcus Vinicius Silva^{***},
Paula Xavier dos Santos^{****}

^{*}bfonseca@cdts.fiocruz.br

Center for Technological Development in Health, Oswaldo Cruz Foundation, Av. Brasil 4036, Rio de Janeiro, 21040-361 (Brazil)

^{**}ricardosampaio@fiocruz.br

Diretoria Regional de Brasília, Oswaldo Cruz Foundation, Av L3 Norte, s/n, Campus Universitário Darcy Ribeiro, Gleba A, Brasília, 70910-900 (Brazil)

^{***}marcussilva@fiocruz.br

Casa de Oswaldo Cruz, Oswaldo Cruz Foundation, Av. Brasil 4036, Rio de Janeiro, 21040-361 (Brazil)

^{****}paula.xavier@fiocruz.br

Vice-Presidency of Education, Information and Communication, Oswaldo Cruz Foundation, Av. Brasil, 4365, Rio de Janeiro, 21040-900 (Brazil).

ABSTRACT

Brazil has been encouraging the establishment of research networks to address strategic health issues in response to social demands, creating an urgent need to develop indicators for their evaluation. The Oswaldo Cruz Foundation (Fiocruz), a national research, training and production institution, has initiated the development of an “Observatory in Science, Technology and Innovation in Health” to monitor and evaluate research and technological development for the formulation of institutional policies. In this context, we are proposing the use of social network analysis to map cooperation in strategic areas of research, identify prominent researchers and support internal research networks. In this preliminary study, co-authorship analysis was used to map the cooperative relations of Fiocruz in tuberculosis (TB) research, an important public health issue for which diagnosis and adequate treatment are still challenging. Our findings suggest that Brazilian research organizations acting in TB research are embedded in highly connected networks. The large number of international organizations present in the Brazilian network reflects the global increase in scientific collaboration and Brazil’s engagement in international collaborative research efforts. Fiocruz frequent cooperation with high-income countries demonstrates its concern in benefiting from the access to facilities, funding, equipment and networks that are often limited in its research setting. Collaboration with high burden countries has to be strengthened, as it could improve access to local knowledge and better understanding of the disease in different endemic contexts. Centrality analysis consolidated information on the importance of Fiocruz in connecting TB research institutions in Brazil. Fiocruz Observatory intends to advance this analysis by looking into the mechanisms of collaboration, identifying priority themes and assessing comparative advantages of the network members, an important contribution to help bridging the translational gap in TB research.

INTRODUCTION

The analysis of science and technology (S&T) networks can provide useful information for research monitoring and evaluation, decision-making processes and for the development of the institutions involved (Fonseca et al., 2016). Collaborative networks for health innovation are particularly important in developing countries where resources are scarce and research capacities are fragmented (Morel, 2005).

Brazil has been encouraging the establishment of research networks to address strategic health issues in response to social demands, creating an urgent need to develop indicators for their evaluation. The Oswaldo Cruz Foundation (Fiocruz), a national research, training and production institution linked to the Ministry of Health, has initiated the development of an “Observatory in Science, Technology and Innovation in Health” to monitor and evaluate research and technological development for the formulation of institutional policies. In this context, we are proposing the use of social network analysis (SNA) to map cooperation in strategic areas of research, identify prominent researchers and support internal research networks. SNA has been previously used to map and measure relationships between researchers and institutions, providing input for research policy (Vanderleist, 2015).

In this preliminary study, co-authorship analysis was used to map the cooperative relations of Fiocruz in tuberculosis (TB) research, an important public health issue for which diagnosis and adequate treatment are still challenging. The paper aims to contribute to the discussion of the following questions: i) How is Fiocruz engaged in the Brazilian TB research networks?; ii) What is its pattern of collaboration?; iii) Can SNA be step stone for supporting S&T organizational management?

METHOD

Scientific articles on TB published by Brazilian organizations were retrieved from the Web of Science database for the period 2005 to 2014. The unit of analysis consisted of S&T institutions where Brazilian-based authors and their national and international collaborators were affiliated at the time of publication. Multiple affiliations were all used in the analysis. As in the health sciences is common to researchers to be affiliated to both university or research institute and hospital or clinic in order to access clinical data and patients (Mattson, 2016), we assumed that these individuals provide a collaboration link between these institutions. In the network, each node is an institution and two institutions were considered connected if its members shared the authorship of a paper.

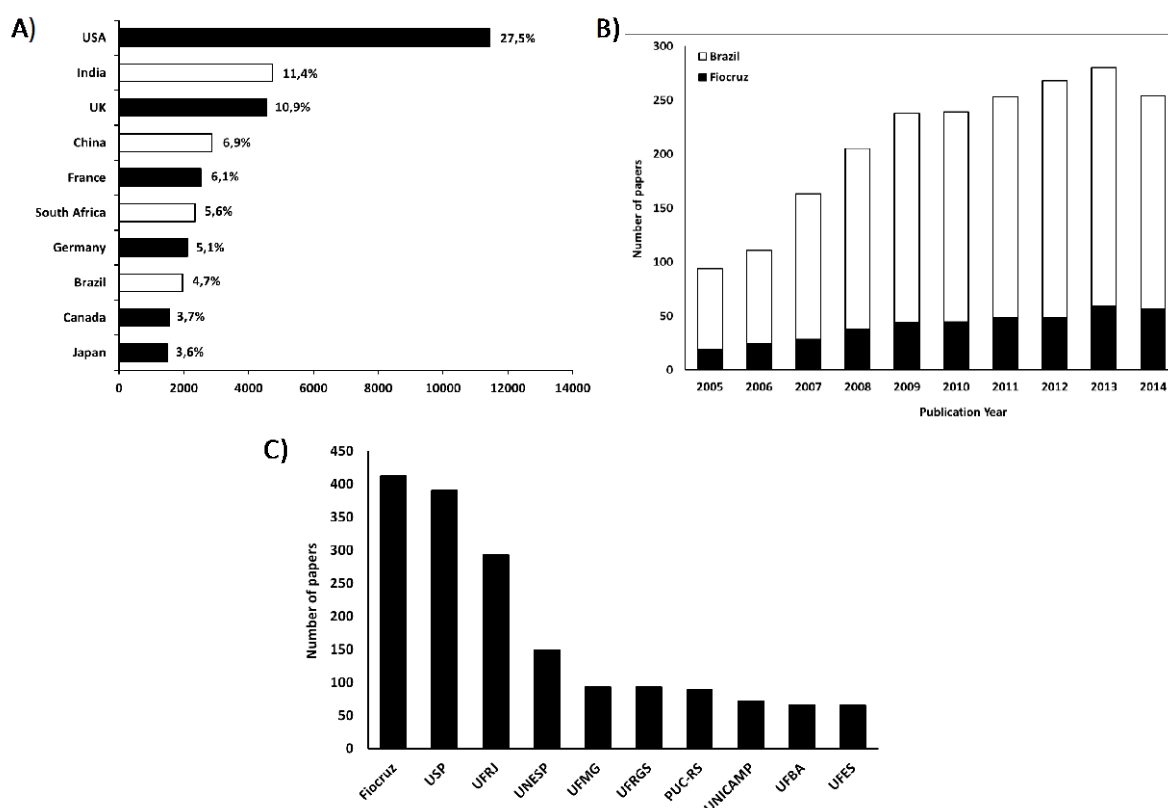
Network connectivity was assessed by the following indicators (Wasserman & Faust, 1994): i) number of nodes and links, corresponding to the number of organizations and their connections; ii) size of the giant component, indicating the maximal subset of fully connected nodes; iii) average degree, indicating the mean number of collaborators the nodes have; iv) average path length, estimating the average smallest number of connections to reach any node in the network; and v) average clustering coefficient, measuring the extent to which the nodes establish a fully connected cluster. Organizations that had prominent roles were identified by their degree centrality - which indicates the number of node’s direct connections -, and by betweenness centrality - which reflects the extent a node acts as a “bridge” between other nodes (Freeman, 1979).

RESULTS

Brazilian organizations accounted for approximately 5% of the world’s scientific production on TB and ranked 4th when compared to other countries with high disease burden (Figure 1A). Fiocruz and Brazil had the same publication trend overall, with a growth until 2009 followed

by a relatively stable production (Figure 1B). Fiocruz was the most productive institution in the country and accounted for approximately 24% of TB publications during the study period (Figure 1C). Although Fiocruz has units in different states in the country, the one located in Rio de Janeiro was responsible for 79% of all Fiocruz publications (n=326), still above the second most productive institution, the University of São Paulo (n=246). Other units involved in TB research were located in the states of Amazonas (2.9%), Bahia (5.3%), Pernambuco (8.2%), Mato Grosso (0.2%), Minas Gerais (4.1%) and Paraná (0.2%). Inter-unit collaboration is modest, with a maximum of five papers in co-authorship between different units.

Figure 1: Tuberculosis research publications by Fiocruz (2005-2014). A) Ten most productive countries and relative contributions (%). White bars indicate high TB burden countries. B) Annual evolution of Brazil's scientific production. B) Top ten most productive Brazilian institutions¹.



The Brazilian organizational network for TB research is composed of 442 national and 670 international institutions from 87 countries (Figure 2).

The United States, United Kingdom and France are Fiocruz most frequent partners, accounting for 18%, 13% and 7%, respectively, of all papers published. Fourteen high disease burden countries have collaborated with Fiocruz, but their association is less frequent. Among

¹ Fiocruz: Oswaldo Cruz Foundation; USP: University of São Paulo; UFRJ: Federal University of Rio de Janeiro; UNESP: São Paulo State University; UFMG: Federal University of Minas Gerais; UFRGS: Federal University of Rio Grande do Sul; PUC-RS: Pontifical Catholic University of Rio Grande do Sul; UNICAMP: State University of Campinas; UFBA: Federal University of Bahia; UFES: Federal University of Espírito Santo.

these countries, South Africa was Fiocruz most frequent partner with 13 papers (3%) in co-authorship.

Figure 2: Institutional network for tuberculosis research involving Brazilian organizations. Nodes are color coded - dark gray for Brazil and light gray for foreign organizations. The size of the nodes is proportionate to their degree centrality. For visualization purposes only the giant component is shown. The top three Brazilian organizations with highest degree centrality are labeled.



The connectivity indicators of the Brazilian TB research network are presented in Table 1

Table 1: Connectivity indicators of TB research networks involving Brazilian organizations.

Indicator	Value
Number of nodes (organizations)	1,112
Number of links	9,138
Giant component size	97.8%
Average degree	16.8
Average clustering coefficient	0.836
Average path length	2.71

The large size of the giant component together with the high average clustering coefficient and average degree, and a low average path length suggest that the network structure is potentially very effective in knowledge generation (high connectedness) and knowledge sharing and diffusion (low distance).

Centrality analysis allowed the identification of the most influential Brazilian organizations (Table 2). Central organizations usually have greater access and control over resources, leading knowledge exchange and preventing many groups from isolation and, in consequence, are more likely to be associated with innovative activities.

Table 2: Top three central organizations of the Brazilian TB research network²

Rank	Organization	Degree centrality	Organization	Betweenness centrality
1	Fiocruz	0,358	Fiocruz	0,249
2	UFRJ	0,260	USP	0,184
3	USP	0,228	UFRJ	0,155

Fiocruz has the highest degree and betweenness centrality values, indicating its prominent role in connecting TB research institutions in Brazil. The Federal University of Rio de Janeiro (UFRJ) and the University of São Paulo (USP) also have important participation in the network.

Fiocruz most frequent national partners are universities, including UFRJ and the Federal University of Pernambuco (UFPE), with 131 (32%) and 32 (7%) papers in collaboration. Frequent international collaborators include the Johns Hopkins University and the University of London.

DISCUSSION

In this preliminary study the evaluation of co-authorship networks identified structural and organizational patterns of TB research involving Brazil and Fiocruz. Our findings suggest that Brazilian research organizations are embedded in highly connected networks.

The large number of international organizations present in the Brazilian network reflects the global increase in scientific collaboration and Brazil's engagement in international collaborative research efforts. Fiocruz frequent cooperation with high-income countries demonstrates its concern in benefiting from the access to good laboratory facilities, funding, equipment and networks that are often limited in its own research setting. Collaboration with high burden countries has to be strengthened, as it could improve access to local knowledge and better understanding of the disease in different endemic contexts.

Centrality analysis consolidated information on the importance of Fiocruz to TB research in Brazil. As a central organization, it has a large number of connections (degree) and is likely to control knowledge flow in the network (betweenness), helping to both disseminate knowledge and facilitate access to resources and research opportunities. Together with UFRJ and USP, they probably had a vital role in maintaining the connection between the overall research

² Fiocruz: Oswaldo Cruz Foundation; USP: University of São Paulo; UFRJ: Federal University of Rio de Janeiro.

network and in ensuring that less well connected or peripheral organizations gained access to new knowledge and information, reducing network vulnerability.

Fiocruz Observatory intends to advance this analysis by looking into the mechanisms of collaboration, identifying priority themes and assessing comparative advantages of the network members, an important contribution to help bridging the translational gap in TB research.

The identification of Fiocruz individual researchers who are most likely to sustain scientific productivity and networking is yet to be evaluated. Such leading authors are expected to be important opinion makers and could assist in guiding the formulation of institutional policies and the promotion of research for public health and development.

REFERENCES

- Fonseca, B.P.F., Fernandes, E. & Fonseca, M.V.A. (2016). Collaboration in science and technology organizations of the public sector: a network perspective. *Science and Public Policy*. In press.
- Freeman, L.C. (1979). Centrality in social networks conceptual clarification. *Social Networks*, 1, 215-39.
- Mattsson, P. (2016). The Evolution of Collaborations in Health Sciences Measured by Co-authorship. *Collaboration Across Health Research and Medical Care: Healthy Collaboration*. Routledge: London.
- Morel, C. M., Acharya, T., Broun, D., Dangi, A., Elias, C., Ganguly, N. K. & Yun, M. (2005). Health innovation networks to help developing countries address neglected diseases. *Science*, 309, 401-4.
- Vanderleist, D. Social Network Analysis as a tool for research policy. (2016). *PLoS Neglected Tropical Diseases*, 9(12):e0004266.
- Wasserman, S. & Faust, K. (1994). *Social Network Analysis: methods and applications*. Cambridge University Press: London.



Partial alphabetical authorship in medical research: an exploratory analysis¹

Philippe Mongeon^{*}, Elise Smith^{*}, Bruno Joyal^{**} and Vincent Larivière^{*,***}

^{*}philippe.mongeon@umontreal.ca; elise.smith@umontreal.ca

École de bibliothéconomie et des sciences de l'information, Université de Montréal, C.P. 6128, succursale Centre-ville Montréal, QC, H3C 3J7 (Canada)

^{**}bruno.joyal@mail.mcgill.ca

Department of Mathematics and Statistics, McGill University, Burnside Hall, Room 1018, 805 Sherbrooke W. Montreal, QC, H3A 0B9 (Canada)

^{***}vincent.lariviere@umontreal.ca

Observatoire des Sciences et des Technologies (OST), Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Université du Québec à Montréal, Montréal (Canada)

INTRODUCTION

Co-authorship has been rising throughout the twentieth Century (Larivière, Sugimoto, Tsou, & Gingras, 2015; Wuchty, Jones, & Uzzi, 2007), to a level that has reached, in certain disciplines, what Cronin (2001) called *hyperauthorship*. With large numbers of scholars contributing to articles comes the challenge of determining the author's order. In medical research, the first(s) and last(s) author positions are attributed to those who made the most important contributions. Typically, the first and last authors are, respectively, junior and senior scholars. Between these two poles are generally listed authors whose contributions are considered less substantial – often lab technicians (Pontille, 2004). In projects involving a large number of researchers, determining the order of authors who made marginal contributions to the research may become particularly difficult. As a result, research teams might choose to order these authors alphabetically, while maintaining a contribution-based order for the others. The trend of partial alphabetical ordering shown by Waltman (2012) seems to support this hypothesis.

However, identifying articles that show traces of such alphabetical order is far from trivial, as one has to distinguish alphabetical order occurring by chance from voluntary alphabetical order (Zuckerman, 1968). Waltman (2012) mitigated this limitation in his analysis the evolution of the frequency of alphabetical ordering, using the probability of chance alphabetical authorship as the margin of error. He also investigated the use of partial alphabetical ordering using an alphabetization score, which is the number of alphabetically ordered consecutive author pairs divided by the total number of consecutive author pairs in the authors list. We take the investigation of partial authorship a step further by identifying and analysing the frequency and the size of subsequences of alphabetically ordered authors. Our study aims to provide answers to the following research questions:

- 1) How prevalent is partial alphabetical order in the medical literature?

¹ This work was supported by the Canada Research Chairs Program and by the Social Sciences and Humanities Research Council of Canada

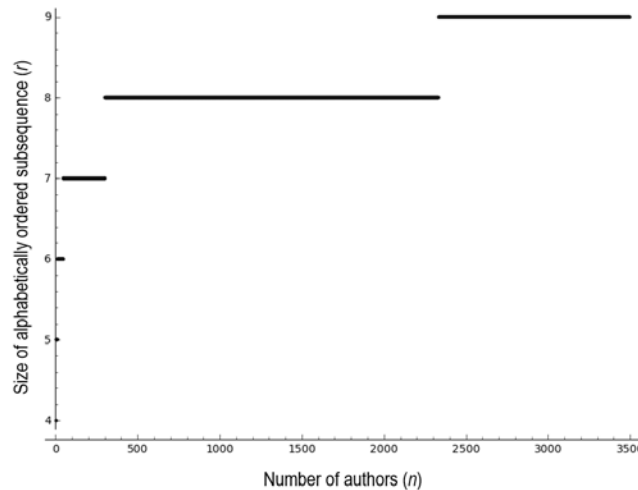
- 2) For articles where an alphabetically ordered subsequence of authors is found, what proportion of authors are listed outside (before or after) that subsequences?
- 3) How does the use of partial alphabetical order (found in questions 1 and 2) evolve over time and as a function of the number of authors per articles?

DATA AND METHODS

Articles in biomedical research and clinical medicine (NSF field classification) with between 6 and 100 authors were retrieved from Thomson Reuters' Web of Science for the 2005-2014 period ($n = 1,967,776$). Like Waltman (2012), we identified in alphabetically ordered pairs of consecutive authors, starting from the first author and moving down the list. An alphabetical subsequence of r authors is formed by consecutive alphabetically ordered pairs. However, the probability of finding an ordered subsequence of r authors by chance increases with the total number (n) of authors. Thus, to mitigate this effect, we calculated a threshold value of r for each value of n , which is the required size of the alphabetical subsequence so that the probability of chance alphabetical order is 5% or lower.

To do so, we define $b_n(r) = n! - s_n(r)$, which is the number of permutations of n authors that do *not* contain an alphabetically ordered subsequence of r authors. One can show by a combinatorial argument (See Elizalde and Noy, 2003) that the exponential generating function

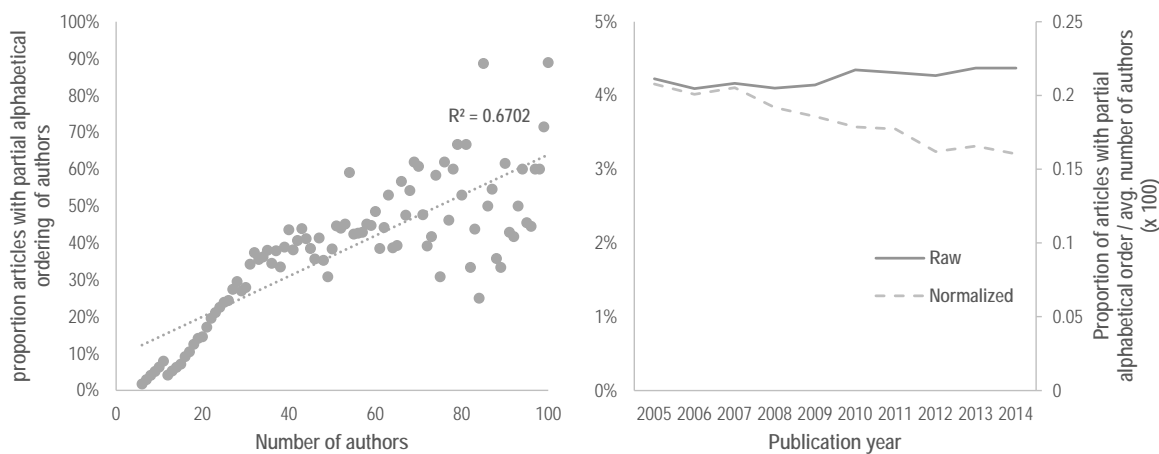
Figure 1: Minimum size r for which the probability of chance alphabetical order is 5% or less as a function of n



RESULTS AND DISCUSSION

Figure 2 (left panel) shows a strong correlation between the number of authors of an article and the frequency of partial alphabetical ordering. While this practice is quite rare for articles with a low number of authors, it becomes more frequent as the number of authors increases.

Figure 2: Prevalence of partial alphabetical ordering of authors, as a function of the number of authors (left panel) and of publication year (right panel)

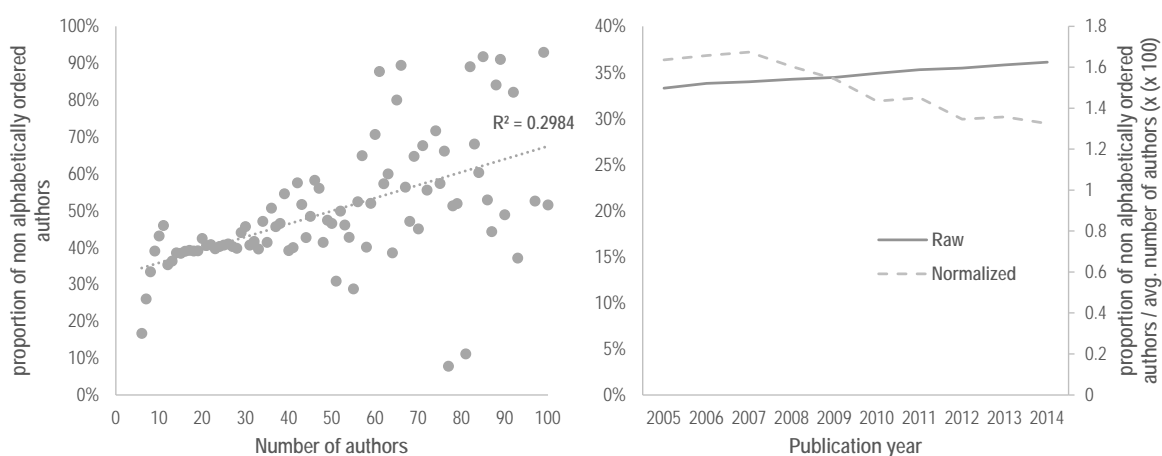


The right panel of Figure 2 shows two opposing trends, depending on the calculation method used. The raw measure indicates that the overall proportion of articles in which partial alphabetical ordering is used tends to grow over time. However, after dividing this proportion by the average number of authors per articles for each year, we see that when the number of authors is fixed, the use of partial alphabetical ordering of authors tends to decrease over time.

For articles where an alphabetically ordered subsequence of authors was found, let us divide the authors in two groups: those listed in alphabetical order—analysed above—and those who are not. While this is an oversimplification of a complex reality, let us also assume that the

authors in the first group are those who made minimal contributions to the work and that the authors in the second group made more substantive contributions. The left panel of figure 3 suggests that the growth in the number of authors is, at least in medical research, mostly due to an increase in substantial contributors. As in figure 2, the right panel of figure 3 shows two opposing trends suggesting that the observed increase in the proportion of substantial contributors over time might be caused by the increasing mean number of authors per articles over time.

Figure 3: Proportion of non-alphabetically ordered authors in articles containing a subsequence of alphabetically ordered authors, as a function of the number of authors (left panel) and of publication year (right panel)



Note: Results are based on the subsample of 80,915 articles for which a *single* alphabetically ordered subsequence was found.

CONCLUSION AND OUTLOOK

These findings provide original insights on the prevalence of partial alphabetical order in medical research and on its evolution over the last decade. We see that as the number of authors per article increases, the prevalence of partial alphabetical order also increases. A plausible explanation is that a high number of authors complexifies their ordering, so only the most important contributors are ordered according to their contribution while the others are ordered alphabetically. Or, since more credit is usually given to the first and last authors, researchers might find that ordering a high number of middle authors by contribution is simply not worth the time.

We also found that a higher proportion of authors are ordered by contribution as the number of authors increases. This suggests that the increase in the average number of authors per article in medical research might be due more to an increase in the number of authors who make substantial contributions rather than an increase of the less substantial contributors being listed as authors. However, our results suggest that when keeping the number of authors fixed, the proportion of authors who made substantial contributions seems to be slightly decreasing over time.

As in Waltman (2012), a limitation of this study is that compound names are concatenated into one string, which might cut some alphabetically ordered subsequences, resulting in a

minor underestimation of the prevalence of partial alphabetical ordering or the subsequences' size. Further developments of this work will address this limitation, and will also provide a broader picture by expanding the period and including other disciplines. This work will provide an enriched understanding of the use of partial alphabetical ordering of authors in science and might have important implications for research evaluation.

REFERENCES

- Elizalde, S., Noy, M. (2003). Consecutive patterns in permutations. *Advances in Applied Mathematics*, 30(1-2), 110-125. Doi:10.1016/S0196-8858(02)00527-4
- Cronin, B. (2001). Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices? *Journal of the American Society for Information Science and Technology*, 52(7), 558–569. doi:10.1002/asi.1097
- Larivière, V., Sugimoto, C.R., Tsou, A., Gingras, Y. (2015). Team Size Matters: Collaboration and scientific impact since 1900. *Journal of the Association for Information Science and Technology*, 66(7): 1323–1332.
- Pontille, D. (2004). *La signature scientifique □: une sociologie pragmatique de l'attribution*. Paris: CNRS.
- Waltman, L. (2012). An empirical analysis of the use of alphabetical authorship in scientific publishing. *Journal of Informetrics*, 6(4), 700–711. doi:10.1016/j.joi.2012.07.008
- Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The Increasing Dominance of Teams in Production of Knowledge. *Science*, 316(5827), 1036–1039. doi:10.1126/science.1136099
- Zuckerman, H. A. (1968). Patterns of Name Ordering Among Authors of Scientific Papers: A Study of Social Symbolism and Its Ambiguity. *American Journal of Sociology*, 74(3), 276–291. doi:10.2307/2775535



The bibliometric behaviour of an expanding specialisation of medical research

Jonathan M Levitt* and Mike Thelwall*

[*J.M.Levitt@wlv.ac.uk](mailto:J.M.Levitt@wlv.ac.uk); m.thelwall@wlv.ac.uk

Statistical Cybermetrics Research Group, University of Wolverhampton, Wulfruna Street, Wolverhampton (UK)

ABSTRACT

This study investigates macular disease research and cataract research, which are both specialisations of Ophthalmology. Macular disease and cataracts are amongst the three leading causes of blindness in the world. Macular research expanded between 1992 and 2006 in that the proportion of Ophthalmology articles classified as macular increased by over 300% in that period. By contrast, during that same period the proportion of Ophthalmology articles classified as ‘cataract’ decreased by over 20%. This study investigates the bibliometric differences between the rapidly expanding specialisation of ‘macular’ and the slightly contracting specialisation of ‘cataract’. Our rationale for investigating these bibliometric differences is that previous researchers have suggested that articles in expanding specialisations are likely to be more highly cited than articles in relatively static specialisations, and it seems important, when comparing specialisations, to try to ensure that articles in a relatively static specialisation are not penalised.

This study first identifies substantial macro-level bibliometric differences between the two specialisations and then gauges the extent to which these differences were associated with the expansion of Macular compared with Cataract. The initial investigation uses coarse-grained delineations of the specialisation, formed from search terms frequently associated with macular (and cataract). It finds that articles in the relatively expanding specialisation were substantially more highly cited and that these differences were associated with the expansion of the specialisation rather than the size of the specialisation (the Matthew effect).

A major limitation of this study is that its coarse-grained delineation of specialisations fails to identify substantial numbers of articles in the specialisation. A more fine-grained delineation using PubMed’s Medical Subject Headings (MESH) has been piloted and additional articles identified. The use of MESH will be investigated further before the conference and our subsequent findings described in our conference presentation.

INTRODUCTION

Over the last decade citation analysis has been increasingly used to compare individuals, departments, universities and countries. These comparisons take into account factors that affect citation, for instance the year of publication and the subject categories of the research.

Although it has been suggested that articles in an expanding specialisation are more likely to be more highly cited than articles in a relatively static specialisation (Peters & van Raan, 1994; Katz, 2000), comparisons do not usually take into account whether a research specialisation is expanding. Without correcting for the growth in specialisations, indicators might unjustifiably shape research agendas by giving too much weight to growing

specialisations. This proposal identifies bibliometric differences between an expanding and a relatively static specialisation and evaluates the extent to which these differences are due to the expansion. It compares the citation behaviour of macular disease, an expanding specialisation, and of cataract, a relatively static specialisation. The specialisations were chosen from medical research, as the guidelines of the UK's 2014 Research Excellence Framework state that all medical panels could take into account citation data. Cataract and macular disease are amongst the three leading causes of blindness in the world.

RELATED RESEARCH AND RESEARCH QUESTIONS

The hypothesis that success engenders success within science, so that successful researchers will tend to get more credit, including citations, was proposed by Merton (1968) as an example of his concept of the Matthew effect. He argued that a consequence was that more successful groups of researchers would attract more funding and new researchers, and consequently would expand. Hence the Matthew effect was hypothesised to lead to a size effect in the sense that larger groups would tend to be more successful. This size effect has been found at the department level (van Raan, 2006) and university level (van Raan, 2008), and in that departments and universities with larger numbers of researchers tended to be highly cited. It has also been found at the individual level (Costas, Bordons, van Leeuwen, and van Raan, 2009), as researchers who published more articles generally received more citations per article. As a result of such findings, one paper has argued that impact indicators should be calculated in a way that does not give an apparently unfair advantage to larger groups of researchers, including for entire countries (Katz, 2000).

The current paper examines specialisations rather than departments, institutions or countries and focuses on citation behaviour changes, as the size of a specialisation changes. Van Raan (1994) wrote “It seems likely, but not inevitable, that articles in an expanding specialisation within a field would tend to be more cited than articles in a static specialisation in the same field.” However, no study has sought to investigate the extent to which citation behaviour of an expanding specialisation differs from that of a more static specialisation. This will be described here as the “growth effect”. Our project partly fills this gap by addressing the following research questions:

1. Were there substantial macro-level bibliometric differences between the two specialisations?
2. To what extent were differences in mean citation primarily due to the size effect or the growth effect?

METHOD AND DATA

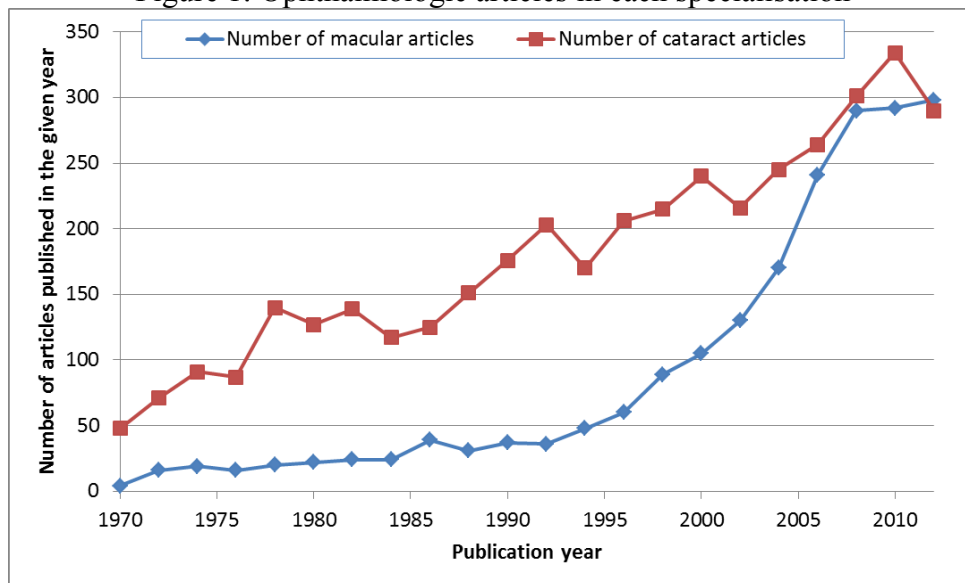
A specialisation is defined as ‘expanding’ over a time period if the ratio, of the number of articles in the specialisation to the number of articles in the parent category, increased over the period; a specialisation is defined as ‘contracting’ over a time period if this ratio decreased over the period. Between 1992 and 2006, Macular was a rapidly expanding specialisation, in that the percentage of Ophthalmology articles classified as Macular increased from 0.83% in 1992 to 3.39% in 2006. Between 1992 and 2006, Cataract was a slowly contracting specialisation, in that the percentage of Ophthalmology articles classified as Cataract decreased from 4.68% in 1992 to 3.72% in 2006.

The two specialisations were categorised on the basis of terms in the titles, for articles in the Web of science (WoS) category of Ophthalmology. The terms categorising macular disease are (a) “macular disease”, “macular diseases”, “macular dystrophy”, “macular dystrophies”, “AMD” and “maculopathy*”, and (b) “age-related” and “macul*”; the term categorising cataract is “cataract*”. The search terms associated with macular disease were derived from fine-tuning the search results of two terms associated with macular disease, ‘macul*’ and ‘AMD’. The term ‘macul*’ was refined to eliminate articles on the macular (a region of the eye) and the search term ‘AMD’ needed to be refined to eliminate articles that used ‘AMD’ to refer to something different from macular disease. The search query for cataract was limited to ‘cataract*’, as cataract is the name of the visual impairment. The search queries for both specialisations are not exhaustive, as many articles on the specialisation may not be identified by our search queries. But this is unlikely to significantly affect the findings.

The size effect was evaluated from the number of articles in the year of publication and the growth effect from the number of articles in the years subsequent to publication.

Figure 1 presents the numbers of Ophthalmologic articles in each specialisation since 1970 (the earliest year for which we have access to WoS data). It indicates much faster expansion between 1992 and 2008 in the macular specialisation than in the cataract specialisation.

Figure 1: Ophthalmologic articles in each specialisation



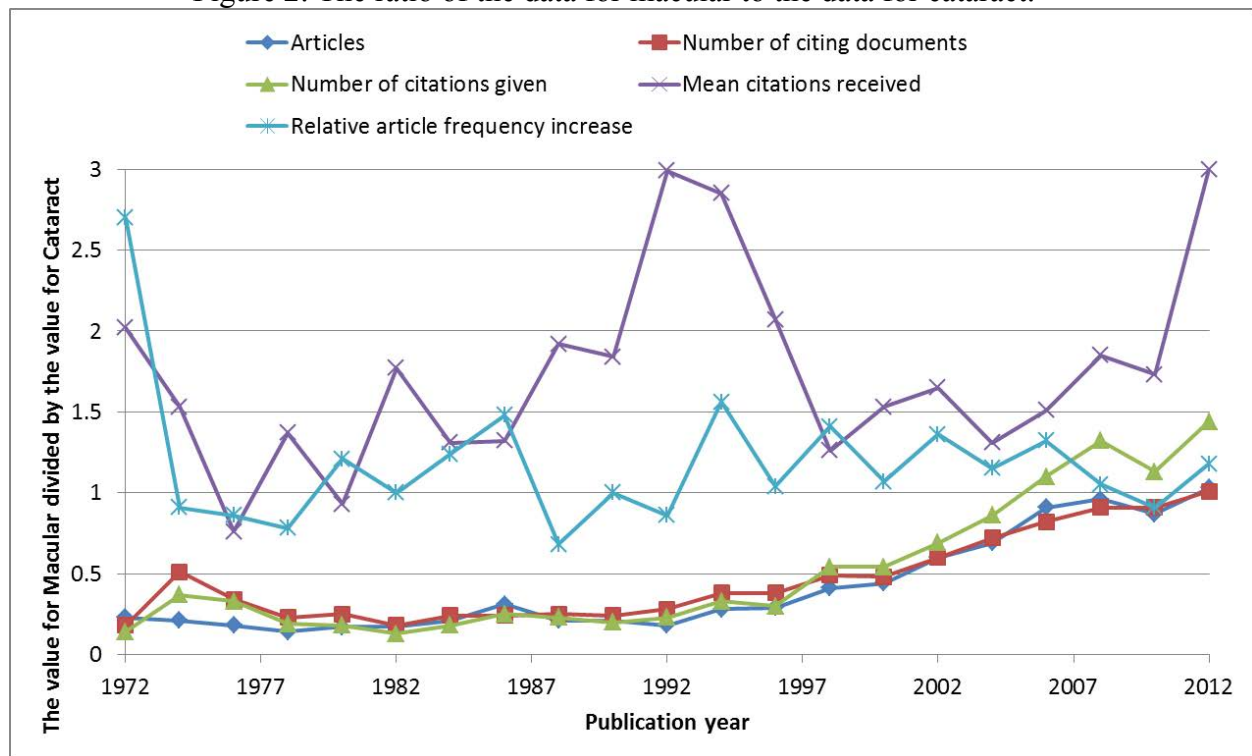
FINDINGS

The results are expressed as ratios between the data for macular to the data for cataract. Figure 2 contains five ratios: ‘Articles’ denotes the macular to cataract ratio of the number of articles in the specialisation published in the year, ‘Citing documents’ the total number of WoS documents in the year citing articles in the specialisation published since 1970, ‘Citations given’ the total number of WoS citations in that year that cite articles in the specialisation published since 1970, and ‘Mean citations received’ the geometric mean of the number of citations received to date by articles in the specialisation published in the year, and ‘Relative article frequency’ denotes the rate of growth (calculated by dividing Articles in the year by Articles in the previous year).

In Figure 2, the number of articles published per year between 1992 and 2006 increased more than five times more rapidly for macular disease compared with cataract (the ratio increased from .18 in 1992 to .91 in 2006). The number of citing documents per year increased between 1992 and 2006 by nearly a factor of three for macular disease compared with cataract (it increased from .18 in 1992 to 0.82 in 2006). For every year since 1982 the mean citation of all articles and for the highest two quartiles was higher for macular disease than for cataract. The rate of expansion was particularly high in the period 1994 to 2006.

In Figure 2, 'Mean citations received' tends to undulate as the size of macular increases relative to cataract. This is not indicative of the size effect. Moreover, although the relative size of macular was much larger for 2006 to 2012 than for 1972 to 1996 (the ratio of sizes ranged from 91% to 103% as opposed to 14% to 31%), the mean citation ratios were comparable (the median ratios were 1.79 and 1.65). The absence of a difference, despite significant changes in size, indicates that there was not a significant size effect.

Figure 2: The ratio of the data for macular to the data for cataract.

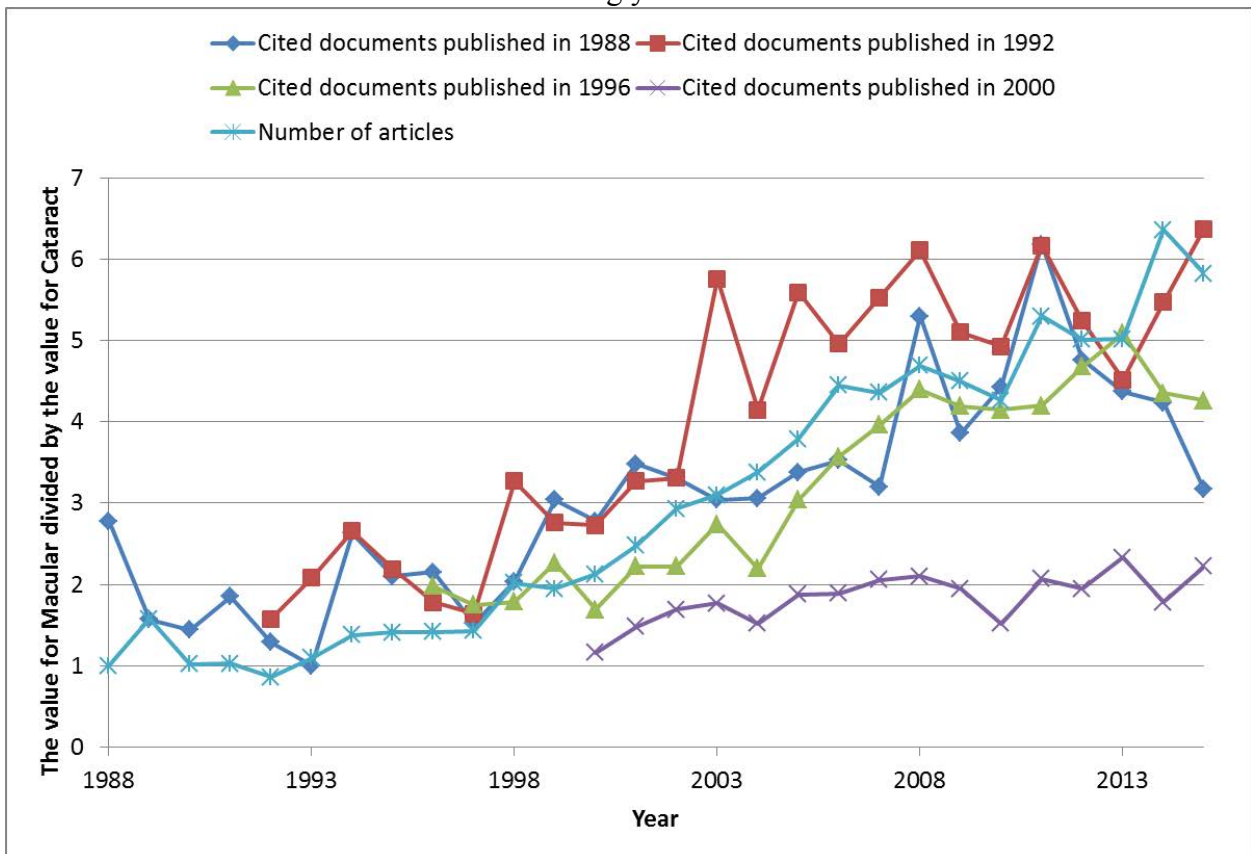


We now test for the growth effect, for which we hypothesise that the number of citations received is roughly proportional to the number of articles published in the year. Our rationale for hypothesising this effect is: (a) we found strong, highly statically significant, Spearman correlations between the number of articles and citing documents (0.96 for macular and .89 for cataract), and (b) it seems reasonable that the number of citations received by an article is roughly proportional to the number of documents citing the specialisation.

In order to test for the cumulative increased citation effect, the annual number of citations received was analysed for all articles published in the years 1988, 1992, 1996 and 2000 and presented in Figure 3. In the figure the ratios of citations received and the ratios of articles in

the year seem to increase roughly in unison. The Spearman correlations between citations received and articles per year were 0.88 for 1988, 0.85 for 1992, 0.90 for 1996 and 0.74 for 2000 (all $p < .01$), tending to support the cumulative increased citation growth hypothesis.

Figure 3: Ratio of Macular to Cataract for citations received in each year since publication. Results are given for four different starting years and normalised by number of articles in the starting year.



Discussion

A reviewer suggested the alternative approach of delineating Macular and Cataract articles, by using search terms obtained from PubMed's Medical Subject Headings (MESH). MESH was used by Skupin et al. (2013) to identify publications within the categories of 'food habits' and 'schools' and by Petersen et al. (2016) to identify publications within the category of 'Diseases'. In this section we briefly investigate how MESH can be used to identify Macular and Cataract articles.

A search on MESH for 'macular' provided 76 subject headings, of which 45 did not contain the term 'macul'. In order to identify the terms that yielded relatively large numbers of macular articles outside the delineation of Macular in this paper, for each of these 45 terms: (a) the set of Ophthalmology articles with the term in the title was isolated, and (b) the subset of this set with 'macular*' in the topic, but neither 'macular*' nor 'AMD' in the title, was isolated. A search on MESH for 'cataract' provided 110 subject headings, of which 20 did not contain the term 'cataract'. In order to identify the terms that yielded relatively large numbers of cataract articles outside the delineation of Cataract in this paper, for each of these 20 terms:

(a) the set of Ophthalmology articles with the term in the title was isolated, and (b) the subset of this set that contain ‘cataract*’ in the topic but not ‘cataract*’ in the title was isolated.

For both macular and cataract, the subsets with particularly high frequencies were identified and the findings presented in Table 1. In the table, the first column presents the term as it appears in MESH, the second column the number of articles in the subset when using the MESH term, the third column presents the variant of the MESH term that seeks also to include associated terms, and the fourth column the number of articles in the subset when using the variant of the MESH term. The cut-off point for inclusion in the table is that the number of extra articles due to inclusion of the MESH term was more than 1% of the number of articles in the Macular and Cataract categories of this study.

Table 1: Frequency of additional articles identified using MESH terms and their variants.

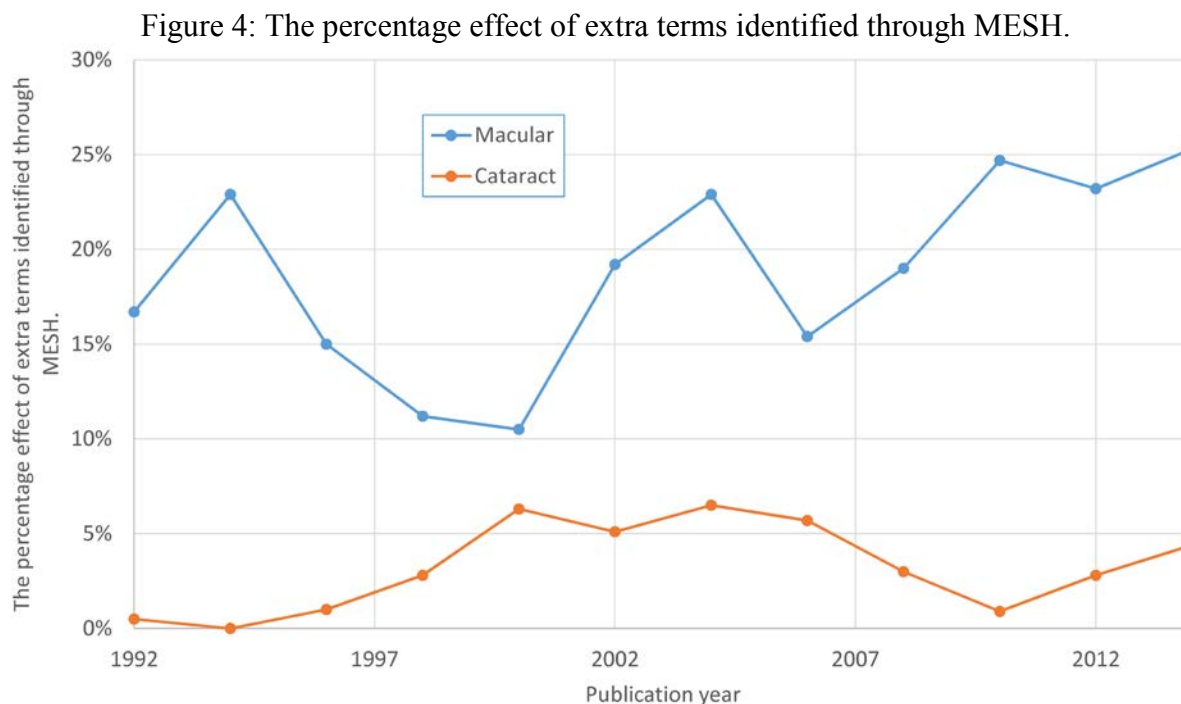
Macular			
MESH term	Articles	Variant of MESH term	Articles
ranibizumab	332	ranibizumab*	332
vitreoretinal surgery	62	vitreoretin*	233
verteporfin	139	verteporfin*	139
retinal degeneration	110	retinal degen*	132
stargardt disease	91	stargardt*	119
geographic atrophy	77	geographic atroph*	77
Cataract			
MESH term	Articles	Variant of MESH term	Articles
Capsule Opacification	228	Capsule Opacificat* OR PCO	231

Table 1 includes six terms associated with macular. Ranibizumab is an anti-angiogenic that has been approved to treat the "wet" type of age-related macular degeneration. Vitreoretinal surgery specialises in treatment of all retina diseases including age-related macular degeneration, retinal detachment, macular holes, and diabetic retinopathy. Verteporfin is a medication used as a photosensitizer for photodynamic therapy to eliminate the abnormal blood vessels in the eye associated with conditions such as the wet form of macular degeneration. Retinal degeneration is the deterioration of the retina (of which the macular is a part) caused by the progressive and eventual death of the cells of the retina. Stargardt disease is an inherited form of juvenile macular degeneration that causes progressive vision loss usually to the point of legal blindness. Geographic atrophy is the advanced form of dry AMD. Table 1 also includes one term associated with cataract. Capsule opacification (PCO abbreviation for Posterior capsule opacification) is a fairly common complication of cataract surgery.

More than four times as many additional articles were identified for macular than for cataract; 1,012 additional articles were identified using variants of the six extra terms associated with macula, whereas only 231 additional articles were identified using variants of the one extra term associated with cataract.

For both macular and cataract, in order to assess the extent to which the number of additional articles varied over time the numbers of additional articles were obtained and for every even numbered year since 1992. The findings are presented in Figure 4. The data in the figure

denote the ratio of the number of additional articles in the year due to the variants of extra terms presented in Table 1 to the number of articles delineated by the criteria used in this paper.



In Figure 4, the effect of including the extra terms on number of articles was substantially higher for macular than for cataract; the effect ranged from 10.5% to 25.3% for macular and from 0% to 6.5% for cataract. Thus, the choice of search method can influence the delineation of specialisations unequally over time. In retrospect, it therefore seems preferable, when bibliometrically investigating medical research, to use more inclusive search criteria that are based on MESH subject headings.

CONCLUSION

A major limitation of this paper is that some of the findings might be affected by the criteria used to delineate the specialisations. Unfortunately, there was insufficient time between receiving the reviews and the resubmission deadline for us to investigate this limitation. However, the limitation will be investigated before the conference and our subsequent findings described in our conference presentation.

Returning to the first research question, there were substantial macro-level differences between the two specialisations, in that: (a) the number of citing documents, and citations given rose more steeply for macular, and (b) macular articles in general received substantially more citations per article.

The second research question asks to what extent were differences in mean citation primarily due to the size effect or the growth effect? Whilst there was no statistically significant evidence for the size effect, there was statistical evidence for the growth effect.

Moreover, the annual number of citations articles received is roughly proportional to the number of articles published in the year. This finding seems unlikely to apply if the specialisations had very different citation patterns (e.g., one was cited much earlier than the other). Because this study is limited to two specialisations, its findings might not apply to other comparisons. It is important when shaping research agendas for indicators to correct for bibliometric differences between specialisations due to expansion. Failure to correct for different rates of growth of specialisations could shape research agendas by giving unmerited funding to growing specialisations.

ACKNOWLEDGEMENT

We wish to thank Gertrude Levitt for her very careful proof-reading.

REFERENCES

- Costas, R., Bordons, M., van Leeuwen, T.N., and van Raan, A.F.J. (2009). Scaling Rules in the Science System: Influence of Field-Specific Citation Characteristics on the impact of Individual Researcher. *Journal of the American Society for Information Science and Technology*, 60(4), 740-753.
- Katz, J. S. (2000). Scale-independent indicators and research evaluation. *Science and Public Policy*, 27(1), 23-36.
- Merton, R.K. (1968) The Matthew Effect in science. *Science*, 159(3810): 56-63.
- van Raan, A.F.J. (2008). Bibliometric statistical properties of the 100 largest European research universities: Prevalent scaling rules in the science system. *Journal of the American Society for Information Science and Technology*, 59(3), 461-475.
- Petersen, A. M., Rotolo, D., and Leydesdorff, L. (2016). A triple helix model of medical innovation: Supply, demand, and technological capabilities in terms of Medical Subject Headings. *Research Policy*, 45(3), 666–681.
- Skupin, A., Biberstine, J. R., and Börner, K. (2013). Visualizing the topical structure of the medical sciences: a self-organizing map approach. *PloS One*, 8(3), e58779.
- van Raan, A.F.J. (2006). Performance-related differences of bibliometric statistical properties of research groups: Cumulative advantages and hierarchically layered networks. *Journal of the American Society for Information Science and Technology*, 57(14), 1919-1935.
- van Raan, A.F.J. (2008). Bibliometric statistical properties of the 100 largest European research universities: Prevalent scaling rules in the science system. *Journal of the American Society for Information Science and Technology*, 59(3), 461-475.

CHAPTER 6

Rethinking Evaluation



Outlining an analytical framework for mapping research evaluation landscapes¹

Fredrik Åström*

*fredrik.astrom@ub.lu.se

Lund University Library, P.O. Box 3, SE 221 00 Lund (Sweden)

ABSTRACT

This paper suggests an infrastructure perspective, as suggested by Star and Bowker (2006), as an analytical framework for studying the research evaluation landscape. An infrastructure is suggested to be understood, not as a concrete technology, but as a system of contextual factors including ‘Actors/Stakeholders’, ‘Technical systems’, and ‘Evaluation practices’. How the framework can be operationalized is exemplified by examples from previous and ongoing research, as well as by identify gaps in current research.

INTRODUCTION

Research evaluation and resource allocation systems permeates academic research, and while evaluation practices per se are well established, there is also a growing literature on research evaluation systems and the effects they are having on ‘the science system’ (de Rijcke et.al, In press).

The aim of this paper is to briefly outline a framework for understanding the complex landscape of research evaluation; and in particular evaluation systems based on the use of bibliometric indicators, to identify from what different perspectives these systems can be analysed and understood as an infrastructure (Star & Bowker, 2006). The basis for developing the framework is examples from previous and current research, as well by identifying gaps in research so far.

BACKGROUND

Over the last three or so decades, we have seen substantial changes in the governance of science (e.g. Whitley and Gläser, 2007); a change that from a policy perspective has been described as change from a linear model to an innovation systems model (e.g. Elzinga, 1995). These changes are often seen as related to the notion of ‘new public management’ (NPM) and the concepts of the audit and/or evaluation society (Dahler-Larsen, 2012).

There have been different suggestions on how we can gain a theoretical understanding of the development of research evaluation systems, both as a general development in research policy and governance, and suggestions of theories contributing to our understanding of particular aspects of the research evaluation systems. There is a long standing discussion in bibliometrics and STS research on the meaning of citations, e.g. drawing on semiotics (Cronin, 2000), or more along the lines of this paper, Wouters’ (2014) suggestion to view the citation as an infrastructure. Recently, Åström and colleagues (2016) suggested ‘boundary

¹ This work was supported by Riksbankens Jubileumsfond: The Swedish Foundation for the Social Sciences and Humanities (SGO14-1153:1)

objects' as a way to theoretically conceptualize scholarly and scientific publications in relation to bibliometrics based research evaluation systems. To understand some of the stakeholders involved in research evaluation processes, Petersohn (In press) has utilized theories on how professions develop. In relation to bibliometrics based research evaluation systems, the conceptualization of research fields and disciplines is also an important aspect, both in terms of how we understand what constitutes fields and disciplines as entities per se (Sugimoto & Weingart, 2015); and how fields are defined in bibliometric analyses and research evaluation systems (Åström et.al, 2016).

Research on the evaluation landscape has been described as having four main research foci: how academic institutions are affected by decreased governmental funding at the same time as NPM related forms of academic governance are introduced, what assessment mechanisms are utilized in national and regional evaluation systems, identifying the dynamics in science and innovation systems, and the effects of indicator use on knowledge production. This last focus address issues of for instance strategic behaviour of scholars/scientists in response to evaluation indicators; and when discussing indicator use in research practices, research on different stakeholders is also brought to attention (de Rijcke et.al, In press).

INFRASTRUCTURES

Star and Bowker (2006) describes infrastructures as representing “one of a number of possible distributions of tasks and properties between hardware, software and people” (Star & Bowker, 2006, p. 232). Drawing on this perspective, we suggest that the evaluation landscape can be understood through the concept of infrastructures, supplying us with an analytical framework for studying evaluation practices. Furthermore, we suggest a categorization of the elements in the evaluation infrastructure in correspondence with Star and Bowker, where “people” take into account the various actors or stakeholders involved in evaluation processes, where “hardware” is understood from the perspective of technical and auxiliary systems, and where “software” represents the evaluation practices per se.

The aspects defined in the categorization are by no means supposed to be considered mutually exclusive, in the same way that categories within these aspects are also often overlapping in many ways. The framework presented here is an attempt at conceptualizing the different aspects of the research evaluation landscape for structured analyses.

“People”: Actors/Stakeholders

The research evaluation landscape is populated by a great variety of actors, such as individual scholars, scientists and research groups; research institutes studying research evaluation; local research administration and services; research funding agencies; national government agencies; research evaluation organizations; and ‘content providers’ (de Rijcke et.al, In press).

There is a variation of types of organizations, from commercial enterprises, over independent research institutes, to public universities and government organizations, all of which taking part in evaluation practices, in academic research on evaluation practices and the formation of research evaluation policies. The roles of these different actors are often intersecting and overlapping; and there is a substantial diffusion of roles and interests both in-between and within groups of actors. The role of university libraries, as part of local research administration and governance, as well as a service institutions for scholars and scientists has been analysed by Åström and Hansson (2013) and Sabrina Petersohn (2016); and Petersohn (Forthcoming) is also studying organizations bordering between being academic research

institutes and research evaluation consultants; and how such expert organizations provide professional expertise for the implementation of national research policy measures.

“Hardware”: Technical & auxiliary systems

The aspect traditionally most associated with infrastructures is technical systems, in the case of bibliometrics based research evaluation, primarily bibliographical databases, citation indices and publication repositories. These exist on many different levels: local, national and international, in terms of coverage, and in terms of where and by whom the databases are developed, from locally developed institutional repositories to international databases produced by large commercial entities. To this can also be added a development where traditional databases are appended by a number of new systems of various kinds: there is a growing market for ‘Current Research Information Systems’ (CRIS), as well as for instance research funding application systems; and to this should also be added systems for bibliometric analyses, where there is a great variation from software developed by individuals to commercial research evaluation tools.

This technical infrastructure has primarily been analysed from perspectives of technical evaluations of the functionality of the systems per se; and the practical applicability of systems in relation to certain evaluation systems and/or practices. Research on the technical infrastructure in a larger context of the research evaluation landscape, however, is rare. This is not for the lack of interesting research questions to address. One issue is of course the implications of – and the different dynamics created by – the use of for instance international citation indices as opposed to locally developed systems. Another complex of questions is related to the increasing communication between systems, where data is being communicated between local publication archives, national research funding application systems, and international citation indices. An example of an attempt at addressing questions related to the technical infrastructures and bibliometrics based research evaluation is recently initiated research on classification issues in relation to bibliometric indicators, where classification systems is seen as a part of a technical infrastructure understood from the point of view of ‘boundary object’ theory (Åström et.al, 2016).

“Software”: Evaluation practices

The part of the research evaluation infrastructure that arguably have received the most attention from scholars and scientists, is the evaluation practices per se. For instance, the relation between national and local resource allocation systems have been investigated in the Swedish context (Hammarfelt et.al, In press), while Hicks (2012) have analysed performance-based university research funding systems from a broader perspective.

An important aspect of the evaluation practices is how they relate to wider research policy issues. The most immediate example is of course resource allocation systems building on publication and/or citation indicators, but equally important is other funding and reward programmes, mandates on issues related to for instance research data management and open access issues.

DISCUSSION

The purpose of this paper has been to suggest an analytical framework for understanding the effects of research evaluation systems on academia and academic research. Aside from studying the effects per se, as in how for instance scholars and scientists adapt to evaluation criteria in their work, a focus on a broad understanding of the infrastructure is presented,

taking into account stakeholders, technical systems and practices. This allows for a structured mapping the evaluation landscape, not the least from a perspective of understanding the ‘materialities’ of research evaluation; and how different aspects of the infrastructure interact.

The complexities found in the evaluation landscape, not the least in terms of how different roles and practices interact, are brought up as an important aspects to consider when analysing regimes of accountability together with the citation infrastructure (Wouters, 2014); strengthening our claim that the infrastructure perspective can be a valuable framework for understanding research evaluation practices as an activity on the borders between science, science policy and research evaluation as a commercial enterprise.

REFERENCES

Åström, F., Hammarfelt, B. & Hansson, J. (2016). Scientific publications as boundary objects: theorizing the intersection of classification and research evaluation. Paper presented the 9th *Conference in the Conceptions of Library and Information Science (CoLIS 9)*, Uppsala, Sweden, June 27, 2016.

Åström, F. & Hansson, J. (2013). How implementation of bibliometric practice affects the role of academic libraries. *Journal of Librarianship and Information Science*, 45(4), 316-322.

Cronin, B. (2000) Semiotics and evaluative bibliometrics. *Journal of Documentation*, 56(4), 440 – 453.

Dahler-Larsen, P. (2012). *The Evaluation Society*. Stanford: Stanford University.

de Rijcke, S., Wouters, P.F., Rushforth, A.D., Franssen, T.P. & Hammarfelt, B. (In press). Evaluation practices and effects of indicator use: A literature review. *Research Evaluation*.

Elzinga, A. (1995). Reflections on Research Evaluation. *Science Studies*, 8(1), 5-23.

Hammarfelt, B, Nelhans, G., Eklund, P. & Åström, F. (In press). The heterogeneous landscape of bibliometric indicators: Evaluating models for allocating resources at Swedish universities. *Research Evaluation*.

Hicks D. (2012). Performance-based university research funding systems. *Research Policy*, 41(2), 251–61.

Petersohn, S. (Forthcoming). Professional evaluation of research excellence: case study of the Netherlands. Paper to be presented at the *4S/EASST Conference*, Barcelona 2016.

Petersohn, S. (2016). Professional competencies and jurisdictional claims in evaluative bibliometrics: The educational mandate of academic librarians. *Education for Information*, 32(2), 165-193.

Star, S.L. & Bowker, G.C. (2006). How to infrastructure. In: L.A. Lievrouw & S. Livingstone (Eds), *Handbook of New Media: Social Shaping and Social Consequences of ICTs* (pp. 230-245. London: Sage.

Sugimoto, C.R & Weingart, S. (2015). The kaleidoscope of disciplinarity. *Journal of Documentation*, 71(4), 775-794

Whitley, R. & Gläser, J. (Eds., 2007). *The Changing governance of the sciences: The advent of research evaluation systems*. Dordrecht: Springer.

Wouters, P. (2014). The citation: From culture to infrastructure. In: B. Cronin & C.R. Sugimoto (Eds), *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact* (pp. 47-66). Cambridge, Mass.: MIT Press

When the Brightest are not the Best

Marco Valente

Università dell'Aquila and LEM, Pisa

marco.valente@univaq.it

Abstract

Selection procedures for new recruits in research organizations, supposedly aiming at identifying the candidates with the highest potential, relies necessarily on indirect information concerning the quality of a researcher. It is safe to assume that this information is correlated to, but not coinciding with, the un-observable future contributions of candidates. There is, consequently, the problem to design a selection procedure such as to ensure the identification of the best candidates.

We show that using exceedingly selective criteria operating on observable proxy indicators of research quality may hinder the overall goal to ensure the highest expected research quality in the hiring organization. The paper presents a simple theoretical model showing that the strategy of pursuing the absolute best during the selection process is very likely to produce worse results than more relaxed selection procedures, aiming humbly at identifying the good candidates.

Keywords: Simulation models, Research assessment, Management of academic institutions.

JEL-classification: A14, H10, C63

1 Introduction

When the demand for resources exceeds the available supply it is necessary to ration some of the potential receivers. Among research institutions the limitation of funds typically generates an excess of candidates in respect of the available positions. Hence, research organizations, such as university departments, need to operate a selection among the candidates on the basis of criteria likely to include, at least to some degree, the assessment of the quality of past research as a proxy for the quality of research expected to be supplied by the new recruit. For this reason, along the increasing diffusion of assessment-based systems to distribute research funds there is a growing attention on the different methods to assess the quality of research and on their overall effects on the institutions involved (Geuna and Martin, 2003; Hicks, 2012).

In this paper we ignore the problems concerning the measurement of research quality and focus on the impact of the selection procedures, based on an idealized indicator and meant to choose, among many candidates, those most likely to provide the best contribution to the research record of the hiring institution. This problem concerns, besides the collection of relevant information, also the issue of the selection procedure to adopt in order to exploit the information. When hiring a new researchers the institutions consider several factors, such as the specialization field, teaching experience, etc. However, a prominent weight in the final decision is increasingly given to the quality of research produced by the candidates as reported by bibliometric indicators, under the assumption that this is an objective measure. However, it is safe to assume that such measures are subject to several potential biases, due to a number of factors. Consequently, the true quality of research cannot be fully identified in the bibliometric indicators, but can, at best, be considered as correlated to the latter.

In this paper we ask whether a competition based solely on the publicly available indicators of research quality is able to select the best candidates, assuming their true, unobservable, quality as researchers is correlated to indicators with some slack, or error. In particular, we wonder whether the increasing the selectivity of the recruiting, and hence the chances to pick those with the highest indicators, is able to also choose those with the highest un-observable quality. We anticipate that this intuitive effect is very small, under fairly general conditions, and that, moreover, adding some additional assumptions we may even reverse the relation between selectivity of the procedure and quality of the recruits.

The next section describes informally an abstract model built to answer the question we posed above. We then provide a formal implementation of the proposed model, implemented in terms of a simple agent-based model. Section 4 presents and discusses the results provided by simulating the model, showing the conditions required for the high and low competitive hiring system to perform better. Before concluding, we sketch briefly an alternative hiring mechanism not relying on a specific estimation of a candidate's quality.

2 Informal model description

The model we propose to investigate the effects of different degrees of competitiveness, described formally in the next section, is a stylized representation of a generic organization, such as a department of a research institution, regularly requiring new recruits to replace retiring members. The model is designed to test the outcome produced by different selection procedures under different external conditions, represented by the nature of the candidates for the positions. The model is meant to highlight a generic properties of a hiring method, and therefore ignores as many details as possible in order to make evident

the relevant consequences for a hypothetical decision maker.

We assume that there exist a correct measure of quality of a researcher¹ that, however, is not directly observable by the modeled decision makers. It is instead possible to collect data (such as, e.g., the publications' record, education record, references, etc.) collectively providing an indicator supposedly approximating the (hidden) quality of the researcher. One of the control parameters of the model is the average gap, assumed stochastic, between the observable indicator available to the simulated agents and the un-observable true quality, that we control in our simulation exercises.

The model represents an ideal "department" that needs to hire new staff in order to replace its retiring members. The new recruit can be chosen only on the base of the available indicators, with the goal is to maximise the average research quality of the department. For simplicity, to highlight the impact of the hiring practices on the quality of an organization, we assume the quality of a hired researcher to remain constant throughout the tenure of selected researchers within the organization, whose length is assumed for simplicity constant and identical for all members. Allowing for the endogenous dynamics of the skills of researchers would complicate the interpretation of the results without adding significantly to the major results, and therefore, at this stage, we prefer to keep this option as a possible future extension.

The model assumes that the hiring method consists in choosing randomly a single new recruit from a set of candidates with probabilities proportional to their observable indicators, so that a "better" candidate is always more likely of being hired than a competitor showing a poorer indicator. The use of a stochastic choice, rather than a deterministic one, reflects the fact that small differences in the indicators' values may have little relevance, similarly to what have been shown to the distributional properties of commonly used research quality indicators, such as the Hirsch's *h-index* (Baccini *et al.*, 2012). The same assumption may be supported by assuming that other considerations enter in the selection procedures besides the research quality indicator. For example, the selecting committee may additionally consider the specific area of specialization of candidates, logistic considerations, reliability, teaching qualities, etc., so that when the indicator of research quality is very similar among two candidates, the probability of selecting one or another is roughly similar, independently from small differences. For modeling purposes we therefore assume that good credentials on research quality provide an advantage for being hired, but that the actual choice is probabilistic. We study a range of different practices differing by the degree of *selectivity* on (observable) research quality, i.e. the relative concentration of the probability distribution assigned to the candidates. A more selective practice assigns far higher probabilities to the (apparently) best candidates, while a more tolerant attitude is expressed by reducing the differences in probabilities, giving therefore relatively more chances to less credited candidates.

As last, and crucial, element of the model we implement a sort of "personal orientation" of researchers. We assume that researchers are aware of the indicators used to assess their work, and that the slack between the publicly available indicators and the true quality of the research is partly under control of the researchers themselves. Yet, pursuing a better public score beyond the "natural" value of a researcher is costly, requiring a diversion from the normal activities of the researcher reflected in an equivalent fall in the un-observable

¹We are assuming that it is possible to represent all the necessarily multidimensional and non-measurable aspects of a researcher's quality by a single value, or even a whole set of values. Moreover, we also assume that the production quality of a researcher remains constant throughout all its career. These assumptions are adopted for obvious reasons of simplifications, and do not undermine our results. On the contrary, they should be interpreted as strengthening them since our most relevant results would be even more evident admitting the non-measurable nature of research quality and the possibility to vary through time.

true quality of the research. We assume that the relation between the voluntarily induced modifications of the observable indicator and true quality are symmetric. That is, a researcher may decide, for example, to invest in ambitious research projects or explore radically new scientific areas causing an increment of her true research quality. This choice, however, implies a fall in the value of the indicator because, for example, the reduction of the number of publications or the resistance of prestigious journals in accepting new ideas challenging the scientific establishment. We represent this feature by assuming the possibility for researchers to devote efforts to either artificially increase their visible indicator of research quality, thus reducing their true quality. On the opposite, a researcher may opt to increase his true research quality, at the cost of worsening the visible indicator. Considering the sensitivity of the assumptions underlining this aspect of the model, the implementation is designed so as to verify its relevance to the eventual results.

3 An agent-based model for researchers' hiring procedures

The model is implemented as a simulation agent-based model using extensively independent random values, so that the average values collected over many time steps and many repetitions ensures a reliable appreciation of the expected results, and of their underlining motivations.

The model contains a group of $N = 100$ "researchers" composing a department, whose generic member is represented by four values:

- $q_i \in [0, 1]$: research quality, unobservable by the agents within the model and whose average over all the members of the department is the main result, measuring the department average research quality; the goal of the selection mechanism is to pursue the highest value of these variables for all the members of the department.
- $i_i \in [0, 1]$: proxy indicator of research quality, observable by the model agents, correlated to the quality q_i as described below.
- $o_i \in [-1, 1]$: preferential orientation ranging from -1 (maximal effort to maximise the public indicator) to 1 (maximal effort to maximise the true research quality).
- $Age_{i,t} \in \{0, 1, 2, \dots, T\}$: age of the researcher, starting from 0 for newly hired members and reaching T , when the researcher retires triggering the department to open a new position.

Besides the age, these values are fixed when the researcher is hired and are not modified during their professional life, assumed to last 100 time steps. At each time step the simulation replaces the retiring members (those reaching $Age_{i,t} = 100$) launching as many independent "calls", each producing eventually one new recruit starting the career in the department with $Age_{i,t} = 0$ ². Each call is answered by exactly 100 candidates whose values are determined as follows, indicating with the modifier x^* the variables for candidate researchers.

The first variable defining a candidate is the true quality of research, q_i^* , drawn from a random value distributed according to a power law. The choice of a skewed distribution reflects the evidence that research skills are distributed in an asymmetric way, with few excellent candidate and increasingly larger numbers of candidates for slowly decreasing

²Note that each open position draws a fresh call, so that the distributional properties affecting each new recruit are independent from the number of calls at each period.

levels of quality. The function adopted produces values for quality levels q distributed according to the distribution $Prob(q = x) = e^{-\alpha x}$, with $x \in [0, 1]$. The higher α the more concentrated the distribution, i.e. smaller the share of top quality values, producing a highly skewed distribution. Conversely lower values for α produce a more even distribution. The extreme value $\alpha = 0$ produces a uniformly distributed random distribution for true research quality spanning evenly over the range $[0 : 1]$.

The raw proxy indicator i_i^* for the candidates is derived from their true quality with a stochastic choice determined by parameter δ . The procedure draws a random value with uniform distribution in the range around the true quality: $i_i^* = U(q_i^* - \delta/2, q_i^* + \delta/2)$. In case the extremes of the range extends beyond the permitted range for researcher's quality $[0,1]$, the range is shifted to ensure that the resulting value is always within unitary interval. That is, if $q_i^* < \delta/2$, then $i_i^* = U(0, \delta/2)$. Symmetrically, for $q_i^* > 1 - \delta/2$ we use $i_i^* = U(1 - \delta/2, 1)$.

The orientation of the candidates is obtained using a uniformly distributed random draw in the range indicated, that is $o_i^* = U(-1, 1)$. After the orientation is determined, the values previously drawn for the quality of research and the proxy indicator are modified as follows: $q_i^* = q_i^* + o_i^* \times \gamma$ and $i_i^* = i_i^* - o_i^* \times \gamma$. Both q_i^* and i_i^* are replaced with the closest boundary if they exceed the range $[0,1]$.

A time step in the simulation run consists in counting the number of members of the department reaching the retirement age and replacing them with new researchers drawn from a set of $N = 100$ candidates, which is recreated at each time step. The probability of choosing a candidate i is computed normalizing the indicators i^* after biasing them with the competitiveness parameter σ : $p_i = \frac{i_i^* \sigma}{\sum_{j=1}^N i_j^* \sigma}$. The parameter σ represents the intensity of the selection process. Higher values of σ represent higher differences in probability, hence favoring candidates with higher indicator values, while lower levels indicate less marked differences in probabilities. Thus, σ may be interpreted as a proxy for the competitiveness of the selection procedure: very high values of σ ensure that almost certainly the candidates with the highest indicator i_j^* will be selected. Lower values of σ instead represent less competitive selections, where candidates with lower indicators will have non negligible chances of being hired, though the probabilities are still proportional to the indicators.

We run a simulation exercise for 10,000 time steps, collecting the average values of the relevant variables across all the time steps. The large sample ensures that the stochastic volatility produced by the random events is absorbed, providing stable results³.

In summary, the model is controlled by the following parameters, whose indicated values will be used for the simulation results presented in the next section:

- $N = 100$, number of, both, members of the department and potential candidates for hiring at any given time step.
- $T = 100$, retirement age for members of the department, after which the researcher is replaced by a new one chosen as indicated among the candidates. Initial age values are chosen randomly, uniformly distributed between 0 and T , to ensure a roughly regular departmental turn over rate.
- $\alpha = 20$, the degree of skewness of the distribution of true quality.

³By stability of results we mean results that do not change by either repeating the simulation run with different series of random values, or by extending the length of simulation runs.

- δ , width of the range for error of the indicators i_i around the true quality q_i , before adjusting for the effects of the orientation of candidates. We set the maximum error at 0.1 on both sides of the true quality, producing a maximum error range of $\delta = 0.2$.
- γ , maximum absolute change induced by the orientation of the researcher in respect of research quality and proxy indicator. The orientation let's the true quality change by a random value drawn uniformly from the interval $[-\delta : \delta]$. We repeat simulation runs for 8 different values of this parameter, ranging from 0 (no effect of orientation) to 0.7 (70% of the maximum quality).
- σ , intensity of selectivity, warping the probabilities of p_i of choosing the candidates. Higher values represent a higher concentration of probability in favor of the candidates with the higher i_i , while lower values represent selection mechanisms proportional to the indicators, but more tolerant for less than stellar values. We consider 10 values from 1 to 10.

4 Simulation results

The simulation described in this section consists in 10,000 time steps during which retiring members of the department are replaced with new recruits chosen from the set of candidates. To minimize the possibility of distortions due to rare random combination of values, the whole set of candidates is fully redrawn at each time step, therefore smoothing away any volatility due to possible extreme values.

The results presented consist in three statistics computed as averages over the researchers hired in the simulated department that, in turn, are again averaged over all the 10,000 time steps. We consider the following average variables measured from the members of the department: average true quality Q ; average proxy indicator I ; average orientation value O . Notice that we have imposed the same random distributions of these variables among candidates. Differences in the results are therefore due to differences among the *selected* candidates, and therefore the distortions between the mean values from the original distributions and the averages computed as results depend solely on the selection procedure. To understand how relevant is the selectivity under different conditions concerning the opportunity of candidates to orient their research attitude, we replicate a simulation run for each combination of the values for the parameters σ (10 values expressing different levels selectivity from 1 to 10), and γ (8 values from 0.0 to 0.7).

We present the results under the setting $\delta = 0.2$, meaning that the maximal difference the proxy indicator and the true quality of a candidate is 0.1, or 10% of the maximum quality, before a possible distortion due to the effects on quality and indicator due to personal orientation.

Figure 1 reports the average indicator values I across all time steps of the researchers hired in the department. The figure shows that increasing the selectivity, giving higher probability to researchers with higher proxy indicators, does indeed increase the average level of this indicator, reflecting the fact that the selectivity works as expected, increasing the probability of hiring candidates with higher indicator values i_i .

The series marked with the value 0, referring the cases in which $\gamma = 0$, can be considered as a sort of benchmark, since the orientation chosen by researchers, in this case, has no effect on neither the true quality nor on the indicator. All other cases are ordered along increasing values of γ , indicating that the stronger the effect of orientation the higher is the average indicator value. This is obvious since when the effect of orientation is stronger (higher γ 's) the higher will be the values i_i^* of the indicators for the candidates,

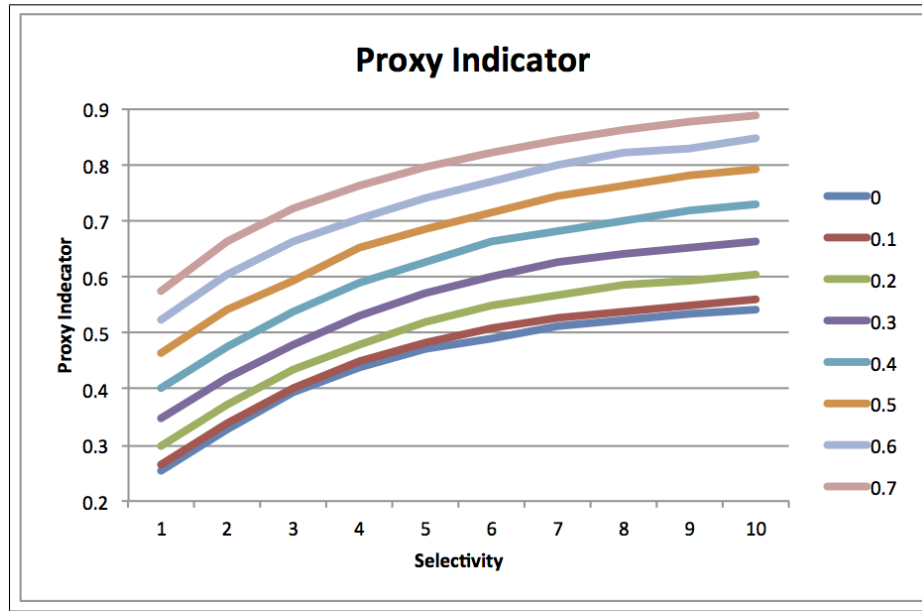


Figure 1: Average proxy indicator I for the quality of researchers for different levels of selectivity levels σ . The different series refer to different values of parameter γ indicating the maximum effect of orientation.

and therefore the selection procedure will have larger pool of (apparently) high quality candidates to choose from.

Figure 2 shows the apparent quality of the department as reported by the proxy indicators paints a far rosier picture than the actual (un-observable) quality of the department as measured by the true quality of its members. This is shown by the fact that the quality levels for all cases are sensibly lower than those reported by the indicators. If the scaling was the only effect, than it still would not matter in terms of the choice of selectivity level (and of a possible incentive policy aiming at influencing personal orientation). But this is not the case.

The ordering of the series for the true qualities is the reverse of the one computed over the indicator. The benchmark case (no effect of orientation) leads the group, while the results produced with the stronger effect of orientation is, by far, the worst. This result is easily explained by the fact that promoting researchers with the best (public) score favors those pushing harder to improve their visible standing, even at the cost of damaging their, invisible, actual research capacity. Remember that the model is built to study the selection process, not behavioural ones. It means that all cases in the same series (same maximum effect of orientation) you have the same distribution of candidates, and therefore the differences depend only on the severity of the selection.

Judging from the true quality, the importance of the selection pressure (reported on the horizontal axis) is actually much reduced, increasing in general quality less than the increment in proxy indicators it produces. Actually, in several cases the average quality provided by even high levels of selectivity remains *below* the expected values from the power law distribution of qualities (about 0.142). This means that a selection committee would do better by picking candidates randomly without any criterion at all, i.e. with uniform probabilities, rather than looking at the proxy indicators of the candidate.

Possibly worse of all from the perspective for a prospective designer of hiring pro-

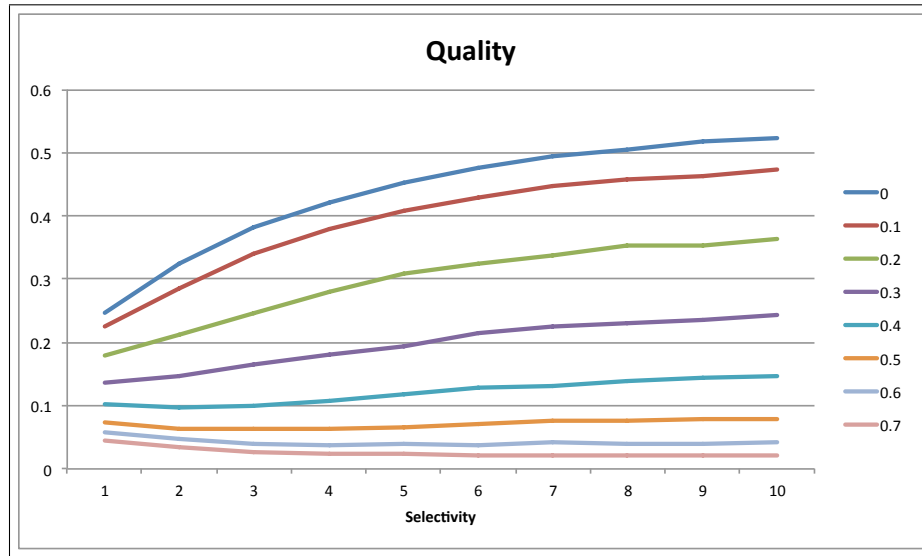


Figure 2: Average (true) quality of researchers for different levels of selectivity levels σ . The series refer to different values of parameter γ indicating the maximum effect of orientation.

cedures, the increment of selectivity is, in some cases, *negatively* related to the quality. Meaning that adopting more stringent selective procedures provides worse results than using a softer intensity of selectivity.

Figure 3 concludes the results showing the average “orientation” of the researchers in the department. We see that the benchmark case shows a null average orientation at any level of selectivity, as can be expected since, when orientation has no effect, we cannot but obtain the expected value of a uniformly distributed random variable in the $[-1,1]$ range, that is 0. For all other cases the results consistently show a strong average negative orientation, meaning that, on average, researchers hired by the department are biased towards improving their public indicators with the consequence of worsening their true research quality. This negative result is accentuated by increasing the level of selectivity, as shown by negative slope of the series; that is, more selecting procedures produce researcher more strongly oriented to massage their public indicator than focusing on research. This result means that selectivity, even when increasing moderately the average quality of hired researchers, does so at the cost of selecting those with the stronger orientation towards focusing on the appearance, rather than substance, of their research.

We can conclude that a recruiting procedure based on selection operated on an indicator of quality systematically overestimates the performance expected from candidates recruited, and frequently fails even to exceed the average performance provided by random choice of candidates. The severity of the selection, represented by the differential in hiring probabilities for candidates with different indicators, is shown to be either poorly and even negatively correlated with the overall average quality for the department, suggesting that hiring criteria should be designed with care to avoid wasting resources (obtaining and elaborating information is costly) to obtain counter-productive results.

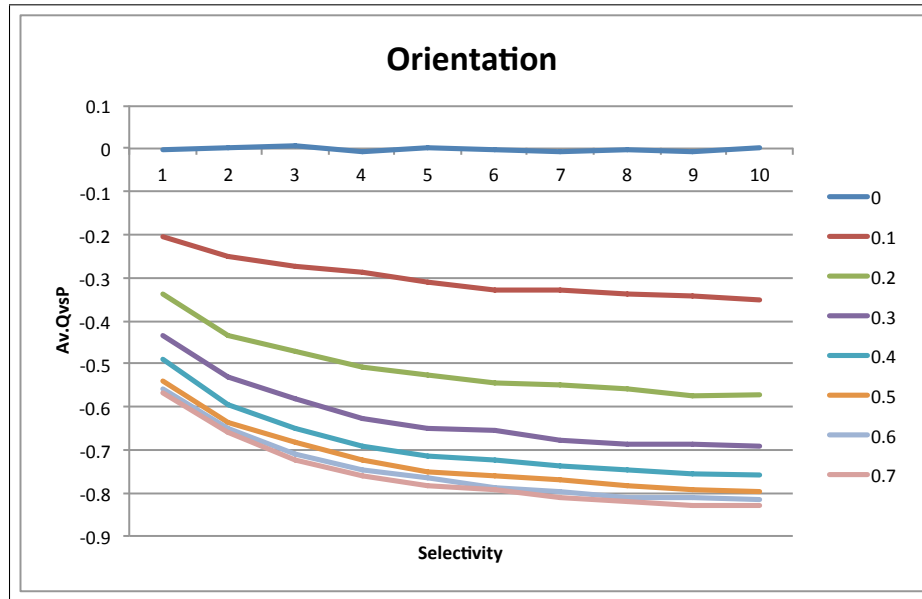


Figure 3: Average orientation between increasing efforts on research (values close to 1) or on improving the proxy indicators (values close to -1). Values produced for different levels of selectivity levels σ . The series refer to different values of parameter γ indicating the maximum effect of orientation.

5 Conclusions

Most of the literature on the assessment of research quality focuses on the problems arising from attaching an estimate to either a researcher or on his/her output. However, the inner mechanisms of the actual procedures using those estimations may be as much, if not more, relevant. This paper has discussed how the selection of new staff may be heavily affected by the selectivity intensity adopted in the hiring procedure, providing counter-intuitive results by means of an agent-based model.

This paper explores the effects of hiring procedures implemented as competitive selection based on the indicators, measuring their performance in terms of the average true quality of the resulting department. We show that, under rather general assumptions, the quality provided by such selection may be pretty poor, even poorer than mere random choice. Moreover, increasing the selection pressure (giving high importance to small differences in indicators) may even lead to *worsen* the performance in terms of average quality of research.

These negative results show that a theoretically perfect system may produce results opposite to the expectations when introducing an apparently minor distortion, in our case that public information is strongly correlated, but not identical, to the true values. In short, a department is better off by not choosing necessarily the (apparently) brightest people it may find, but needs searching more sophisticated hiring systems, more robust against the biases induced by poor information.

References

- ALBERTO, B. and GIUSEPPE, D. N. (2015), “Do they agree? Bibliometric evaluation vs informed peer review in the Italian research assessment exercise”, *arXiv*, <http://arxiv.org/abs/1505.00115>.
- BACCINI, A., BARABESI, L., MARCHESI, M. and PRATELLI, L. (2012), “Statistical inference on the h-index with an application to top-scientist performance”, *Journal of Informetrics*, **6**(4), pp. 721–728.
- BERTOCCHI, G., GAMBARDELLA, A., JAPPELLI, T., NAPPI, C. and PERACCHI, F. (2015), “Bibliometric evaluation vs. informed peer review: Evidence from Italy”, *Research Policy*, **44**, pp. 451–466.
- GEUNA, A. and MARTIN, B. R. (2003), “University Research Evaluation and Funding: An International Comparison”, *Minerva*, **41**(4), pp. 277–304.
- GILLIES, D. (2014), “Selecting applications for funding: why random choice is better than peer review”, *RT. A Journal on Research Policy and Evaluation*, **2**(1).
- GLÄNZEL, W. (2010), “On reliability and robustness of scientometrics indicators based on stochastic models. An evidence-based opinion paper”, *Journal of Informetrics*, **4**(3), pp. 313–319.
- HICKS, D. (2012), “Performance-based university research funding systems”, *Research Policy*, **41**(2), pp. 251–262.
- IOANNIDIS, J. (2011), “More time for research: Fund people not projects”, *Nature*, **477**, pp. 529–531.
- MARTIN, B. R. (2011), “The Research Excellence Framework and the ‘impact agenda’: are we creating a Frankenstein monster?”, *Research Evaluation*, **20**(3), pp. 247–254.



The reward (eco)system of science: More than the sum of its parts?¹

Nadine Desrochers*, Stefanie Haustein*, Juan Pablo Alperin**, Timothy D. Bowman***, Adrián A. Díaz-Faes****, Vincent Larivière*, Philippe Mongeon*, Adèle Paul-Hus*, Anabel Quan-Haase*****, Elise Smith*, Cassidy R. Sugimoto*****

*nadine.desrochers@umontreal.ca; stefanie.haustein@umontreal.ca; vincent.lariviere@umontreal.ca; philippe.mongeon@umontreal.ca; adele.paul-hus@umontreal.ca; elise.smith@umontreal.ca

École de bibliothéconomie et des sciences de l'information, Université de Montréal,
C.P. 6128, Succ. Centre-Ville, Montréal, Québec, Canada H3C 3J7

** juan@alperin.ca

Canadian Institute for Studies in Publishing, Simon Fraser University,
515 West Hastings Street, Vancouver, British Columbia, Canada V6B 5K3

*** timothy.bowman@utu.fi

Sociology of Education Research Center, RUSE
University of Turku, Finland 20014

**** adrian.arias@cchs.csic.es

IFS, Spanish National Research Council (CSIC), Albasanz 26-28, 28037 Madrid, Spain

***** aquan@uwo.ca

Faculty of Information and Media Studies/Department of Sociology, Western University,
North Campus Building, London ON, Canada N6A 5B7

***** sugimoto@indiana.edu

Department of Computer Science and Informatics, School of Informatics and Computing, Indiana University,
919 E. 10th Street, Bloomington, IN, USA 47408

ABSTRACT

In the 1990s, Blaise Cronin and his collaborators established a “reward triangle” of science, which consisted of authorship, citations, and acknowledgements. However, in the last decades, the landscape of scholarly communication and recognition has changed immensely. The use of social media in scholarly communication has generated a new set of indicators, dubbed altmetrics. The proliferation of indicators and the strong tendency to rely on quantitative measures is a fait accompli. Needless to say, the ubiquity of measurement and evaluation creates new forms of pressure and affects scientific behavior. This event will be presented as a twist on an open fishbowl. Out of five chairs, four will be occupied by a first set of participants and each member will represent one of four elements of the reward system of science to open the discussion: authorship, citations, acknowledgements, and social media. A fifth chair will be left empty for audience members to participate. To involve remote audiences, the fishbowl will be live-tweeted. The Twitter feed will be displayed and used as a backchannel. Audience members who may not wish to take a chair will also be able to

¹ This research was supported by the Social Sciences and Humanities Research Council of Canada (Insight Development Grant # 430-2014-0617) and the Alfred P. Sloan Foundation (Grant # G-2014-3-25). The organizers further thank the authors of the ASIS&T panel (Desrochers et al., 2015) and Rodrigo Costas in particular for the discussions around the idea of an “ecosystem” in science.

participate via Twitter. One of the organizers will monitor the Twitter feed, taking the empty chair to relay what Twitter users are saying, in order to address the feed and further the exchanges. During the final 15 minutes of the fishbowl, participants will be asked to come forth and make recommendations pertaining to the initial target topics and any emerging topics. They will be asked to formulate these in short sentences, so that they can be relayed on Twitter; the recommendations will also be made available for further use as one document on etherpad (<http://etherpad.org/>).

PURPOSE AND INTENDED AUDIENCE

In the 1990s, Blaise Cronin and his collaborators established a “reward triangle” of science (Cronin and Weaver-Wozniak, 1993), which consisted of authorship, citations, and acknowledgements. This can be seen as a take on the traditional pillars of scientific endeavors: to research, discover, and disseminate; to have some impact on further research and society; to account for the collaboration, supervision, assistance, review, and infrastructure without which the system would crumble.

However, in the last decades, the landscape of scholarly communication and recognition has changed immensely. Hyperauthorship has boomed (Cronin, 2001), blurring the lines of scientific accountability (Wray, 2006; Mongeon and Larivière, 2014), and causing certain journals to ask for contributorship statements (PLOS, n.d.). The notion of credit has also been brought forth to solve the oft-debated question of who should get authorship or a simple “thank you” (Birnholtz, 2006; Cronin, 1991; Ngai, Gold, Gill and Rochon, 2005; Rennie, Yank and Emmanuel, 1997). Along the same lines, 45 years of acknowledgements research has not yielded clear guidelines on the value and role of this paratext (Genette, 1997) in scientific evaluations (Desrochers, Paul-Hus and Pecoskie, 2015). Citations and derived indicators, which were once the exclusive domain of proprietary citation databases such as the Web of Science and Scopus, are now being ubiquitously displayed by Google and on various publishers’ websites. Originally thought of as “World Brain” and created as a retrieval tool for scholarly literature (Garfield, 1955, 1964), citation indexing has played a major role in transforming the scholarly community into an evaluation society (Dahler-Larsen, 2012). Trends such as the need for swift rises in publication and citation counts have had many pervasive effects, such as the demise of national languages in science (Desrochers and Larivière, in press). The use of social media in scholarly communication has generated a new set of indicators, dubbed altmetrics, that comprise a range of metrics based on online events that were not recorded in previous times (Priem 2014; Haustein, Sugimoto and Larivière, 2015); however, their heterogeneity has led to discussions regarding their meaning and their validity as symbolic capital (Haustein, Bowman and Costas, 2016). Moreover, social media use has introduced boundary issues between the scientific field and the personal realm (Bowman, 2015).

The proliferation of indicators and the strong tendency to rely on quantitative measures is a *fait accompli*. Needless to say, the ubiquity of measurement and evaluation creates new forms of pressure and affects scientific behavior. However, more and more voices are rising within the ranks of the bibliometric and altmetric communities to warn all agents against the Pandora effect the toolbox we are building may unleash. The calls for contextualization made by the Leiden Manifesto (Hicks, Wouters, Waltman, de Rijcke and Rafols, 2015) echo analyses showing the necessary complementarity of quantitative and qualitative measures for the

understanding and adequate evaluation of research (Traag and Franssen, 2016); and the parting editorial of Blaise Cronin (2015) for the *Journal of the Association for Information Science and Technology* revealed only too well some of the effects the pressures of the current system have on the production and dissemination of science.

In order to discuss the various aspects of scholarly communication today, Desrochers et al. (2015) espoused a conceptual framework based on the ideals put forth by Merton's work on the reward system of science (1973) along with Bourdieu's (1975; 1996) conception of academia as a highly codified field, and identified four main elements of the scientific reward system: authorship, citations, acknowledgements, and social media. This brought together the concepts of recognition and symbolic capital, fundamental to the work of the scientific community and its *illusio*, which defines the rules and stakes of the field (Bourdieu, 1996).

The proposed fishbowl is an attempt to further the discussion on how the various elements of the reward system of science combine to create an ecosystem where the symbolic capital associated with each act is understood and valued. Thanks to audience participation, the hope is that it will bring together varied conceptual viewpoints, whether Bourdieusian, Mertonian, Latourian, or other. These can help build critical outlooks on the indicators related to science, its rewards, and the relation of individual indicators to the whole system. Therefore, all agents of academia can—and perhaps should—be part of this discussion; this fishbowl would be a step towards engaging them from the perspective of understanding what the various indicators create as an ecosystem of science.

EVENT FORMAT AND PROCESS

Set up

The event will be presented as a twist on an open fishbowl. Out of five chairs, four will be initially occupied by a first set of participants (“Team 1”) and each member will represent one of the four elements of the reward system of science identified in Desrochers et al. (2015) to open the discussion: authorship (Elise Smith), citations (Cassidy R. Sugimoto), acknowledgements (Adèle Paul-Hus), and social media (Juan Pablo Alperin). A fifth chair will be left empty for audience members to participate.

To involve remote audiences, the fishbowl will be live-tweeted. The Twitter feed will be displayed and used as a backchannel. Audience members who do not wish to take a chair will also be able to participate via Twitter. One of the organizers will monitor the Twitter feed, at times taking the empty chair to relay what Twitter users are saying, in order to further the exchanges.

Introduction

The moderator (Nadine Desrochers) will give a 10-minute introduction to present the format of the fishbowl and briefly introduce the four elements of the reward system.

Position statements (Team 1)

The moderator will then invite each member of Team 1 to give a 3-5 minute statement on their target topic, the role and place of related acts in the system, the indicators that measure the presence and impact of these acts, and their relationship with the other topics and elements of the reward system of science.

Discussion and position statements (team 2)

After this, the fishbowl will open to the audience. Each attendee will be free to take the empty chair, which will then push a member of Team 1 to retire to the audience. Members of Team 2 (Adrián A. Díaz-Faes, Timothy D. Bowman, Stefanie Haustein, Vincent Larivière, Philippe Mongeon, Anabel Quan-Haase) will be waiting to give their statements during any lull in the discussion, allowing for a dynamic mix of prepared and improvised interventions. The moderator will ask all participants to keep their contributions to 3-5 minutes.

Recommendations

During the final 15 minutes of the fishbowl, participants will be asked to come forth and make recommendations pertaining to the initial target elements and any emerging topics. They will be asked to formulate these in short sentences, so that they can be relayed on Twitter; the recommendations will also be made available for further use as one document on etherpad (<http://etherpad.org/>).

Participants will be guided by asking to focus on three points:

1. The scientific activity and its place (or not) in the reward system of science (e.g.: Should acknowledgements be taken into consideration in evaluations?)
2. The relationship between different scientific activities (e.g.: What is more valuable, a patent or a reader count? 10 papers co-signed with 5 authors, or 2 papers as single author?)
3. The relationship between scientific activities and the indicators used to measure them (e.g.: Many indicators are based on the Web of Science and other hard sciences datasets; what are the effects of this on our perceptions and evaluations of research in the Arts & Humanities?)

Rundown of timing is as follows, with maximum times indicated:

- Introduction - 10 minutes
- Team 1 statements - maximum 20 minutes
- Discussion, including Team 2 statements and Twitter-based discussion - 45 minutes
- Final recommendations - 15 minutes

RELEVANCE TO THE CONFERENCE

The discussion points, both for the prepared statements and as put forward by the moderator throughout the event, will touch upon the following aspects of indicators and their uses, of interest to the STI-ENID community:

- Disciplinary differences and practices
- Existing and changing relationships between quantitative indicators and the qualitative nature of peer review
- Social media presence, acts, and policies
- Biases in underlying data
- Challenges in data quality
- Misuses and adverse effects such as misreadings, gaming, or behaviors catering to the indicators in place

However, the overarching questions are the following ones, self-evident perhaps, yet overdue:

- How do—and how can—the various indicators we use to measure, discuss, and evaluate the value and impact of research come together? In other words,

- What is the current ecosystem of science?
- How do its various forms of symbolic capital coexist? And,
- How do these forms of symbolic capital shape research agendas?

NOVELTY

By bringing all aspects of the reward system of science together in one discussion space, by inviting people in and outside the room to engage through speech or social media, this fishbowl should create a forum that looks beyond the singularity of indicators. Colliding and complementary perspectives can then emerge and we can see (and perhaps confront) the perceptions of the scientific field we are currently shaping. The use of Twitter and etherpad will make the event interactive beyond the walls and duration of the conference, as well as create an outcome for further use.

Preferred length: 60-90 minutes, depending on how dynamic the discussion becomes

Preferred number of participants/attendees: 30-40

Special requests/equipment needs: A projector and screen for the introductory PowerPoint slides, Twitter-feed, and etherpad display. The room does not need to be set up in a round per se, but it should allow for up to 6 people to be seated and seen at once, preferably not at a desk, and it should be easy for people from the audience to leave their seat and come to the presentation area.

THE ORGANIZERS/PARTICIPANTS:

Nadine Desrochers holds degrees from Western University and the University of Ottawa. She is an assistant professor at the École de bibliothéconomie et des sciences de l'information, Université de Montréal. Her research on scholarly communication focuses on the reward system of science, information behavior, and paratextuality.

Stefanie Haustein is a postdoctoral researcher at the Université de Montréal. She holds a Master's degree in History, American linguistics and Literature and Information Science and a PhD in information science from Heinrich Heine University Düsseldorf. Her current research focuses on altmetrics, bibliometrics, and scholarly communication.

Juan Pablo Alperin is an assistant professor at Simon Fraser University, Canada. He holds a Master's degree in Geography from the University of Waterloo, and a PhD in Education from Stanford University. He is a multi-disciplinary scholar who uses computational techniques, surveys, and interviews to investigate ways of raising the scientific quality, global impact, and public use of scholarly work.

Timothy D. Bowman holds a Ph.D. in Information Science from Indiana University, Bloomington. His interests include scholarly communication, data science, altmetrics, and web programming. Currently he is a postdoctoral researcher at the University of Turku, Finland, examining large amount of data to determine how scholars communicate and represent themselves online.

Adrián A. Díaz-Faes is a member of the ACUTE group at the Spanish National Research Council (CSIC). He holds a Master's degree and a PhD in Multivariate Applied Statistics from the University of Salamanca. His research focuses on novel approaches to the study of collaboration in science and the development of new bibliometric indicators.

Vincent Larivière holds the Canada Research Chair on the Transformations of Scholarly Communication at the Université de Montréal. He is also the scientific director of the *Érudit* journal platform, and associate scientific director of the Observatoire des sciences et des technologies (CIRST-UQAM).

Philippe Mongeon is a doctoral candidate in Information Science at the École de bibliothéconomie et des sciences de l'information, Université de Montréal. His research focuses on the reward system of science, and more specifically on authorship and inventorship attribution practices in research teams.

Adèle Paul-Hus is a doctoral candidate in Information Science at the École de bibliothéconomie et des sciences de l'information, Université de Montréal. She holds Master's degrees in Information Science and Anthropology from the Université de Montréal. Her research focuses on acknowledgement functions within the scholarly communication system.

Anabel Quan-Haase is an associate professor of Information and Media Studies and Sociology, Western University. She is the author of *Technology and Society: Social Networks, Power and Inequality* (Oxford University Press, 2013; 2nd ed. 2016). Her research focuses on social media, social networks, social capital, and innovation.

Elise Smith recently completed a PhD in the Applied Social Sciences Programme (Bioethics Option) at the Université de Montréal. Her doctoral research focuses on the fair distribution of authorship in multidisciplinary collaborations. She is broadly interested in publication ethics and bias in research linked especially to gender, power, and seniority.

Cassidy R. Sugimoto is an associate professor in Informatics at Indiana University's School of Informatics and Computing. She is the president of the International Society for Scientometrics and Informetrics. Her research focuses on the formal and informal ways in which scholars communicate to each other and to the broader public.

REFERENCES

Birnholtz, J. P. (2006). What does it mean to be an author? The intersection of credit, contribution, and collaboration in science. *Journal of the American Society for Information Science and Technology*, 57(13), 1758–1770. doi: 10.1002/asi.20380

Bourdieu, P. (1975). The specificity of the scientific field and the social conditions of the progress of reason. *Social Science Information* 14(6), 19-47.

Bourdieu, P. (1996). *The rules of art: Genesis and structure of the literary field*. Stanford, CA: Stanford University Press.

Bowman, T.D. (2015). Differences in personal and professional tweets of scholars. *Aslib Journal of Information Management*, 67(3), 356-371.

Cronin, B. (1991). Let the credits roll: a preliminary examination of the role played by mentors and trusted assessors in disciplinary formation. *Journal of Documentation*, 47(3).

Cronin, B. (2001). Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices? *Journal of the American Society for Information Science and Technology*, 52(7), 558–569.

Cronin, B. (2015). The writing on the wall. *Journal of the Association for Information Science and Technology*, 66(5), 873–875.

Cronin, B., & Weaver-Wozniak, S. (1993). Online access to acknowledgements. In *Proceedings of the 14th National Online Meeting 1993* (p. 93-98). New York: M.E. Williams.

Dahler-Larsen, P. (2012). *The evaluation society*. Stanford, CA: Stanford University Press.

Desrochers, N. & Larivière, V. (In press). *Recognition* ou reconnaissance: de la question des langues en diffusion des connaissances. In Borg, S., Cheggour, M., Desrochers, N., Gajo, L., Larivière, V., and Vlad, M. (Eds.). *L'Université en contexte plurilingue dans la dynamique numérique: Actes du congrès annuel de l'Agence universitaire de la Francophonie, Marrakech, November 12-13, 2015*. Paris: Éditions des Archives Contemporaines.

Desrochers, N., Paul-Hus, A., Bowman, T., Costas, R., Haustein, S., Larivière, V., Mongeon, P., Pecoskie, J., Quan-Haase, A., & Tsou, A. (2015). Authorship, patents, citations, acknowledgments, tweets, reader counts and the multifaceted reward system of science. In *Proceedings of the American Society for Information Science and Technology*. <https://www.asist.org/files/meetings/am15/proceedings/openpage15.html>

Desrochers, N., Paul-Hus, A., & Pecoskie, J. (2015). Founding concepts and foundational work: Establishing the framework for the use of acknowledgments as indicators. *Proceedings of the 15th International Conference on Scientometrics and Informetrics*, International Society for Scientometrics and Informetrics (ISSI), Istanbul, Turkey, June 29-July 4, 2015, 890-894. <http://www.issi2015.org/files/downloads/all-papers/0890.pdf>

Garfield, E. (1955). Citation indexes for science. A new dimension in documentation through association of ideas, *Science*, 122(3159), 108–111.

Garfield, E. (1964), 'Science Citation Index'—A new dimension in indexing, *Science*, 144(3619), 649–654.

Genette, G. (1997). *Paratexts: Thresholds of interpretation* (Jane E. Lewin, trans.). Cambridge: Cambridge University Press.

Haustein, S., Bowman, T. D., & Costas, R. (2016). Interpreting “altmetrics”: viewing acts on social media through the lens of citation and social theories. In C. R. Sugimoto (Ed.), *Theories of informetrics and scholarly communication. A festschrift in honor of Blaise Cronin* (pp. 372–405). Boston, MA: De Gruyter.

Haustein, S., Sugimoto, C.R., & Larivière, V. (2015). Social media in scholarly communication. *Aslib Journal of Information Management*, 67(3), 1-14.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden manifesto for research metrics. *Nature*, 520, 429-431.

Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. Chicago, IL: Chicago University Press.

Mongeon, P., & Larivière, V. (2014). The consequences of retractions for co-authors: Scientific fraud and error in biomedicine. *STI 2014 Leiden*, 404.

Ngai, S., Gold, J. L., Gill, S. S., & Rochon, P. A. (2005). Haunted manuscripts: Ghost authorship in the medical literature. *Accountability in Research*, 12(2), 103–14.

PLOS Journals (n.d.). Authorship. Retrieved from <http://journals.plos.org/plosone/s/ethical-publishing-practice#loc-authorship>

Priem, J. (2014). Altmetrics. In B. Cronin & C. R. Sugimoto (Eds.), *Beyond bibliometrics: Harnessing multidimensional indicators of performance* (pp. 263–287). Cambridge, MA: MIT Press.

Rennie, D., Yank, V., & Emanuel, L. (1997). When authorship fails: A proposal to make contributors accountable. *Jama*, 278(7), 579-585.

Traag, V. & Franssen, T. (2016). Revealing the quantitative-qualitative divide in sociology using bibliometric visualization [blog post]. <http://www.cwts.nl/blog?article=n-q2v294&title=revealing-the-quantitative-qualitative-divide-in-sociology-using-bibliometric-visualization#authors>

Wray, K. B. (2006). Scientific authorship in the age of collaborative research. *Studies in History and Philosophy of Science Part A*, 37(3), 505–514.



Capturing knowledge integration through collaborations: measures of the diversity and coherence in multiple proximity dimensions

By Frédérique Lang*, Ismael Ràfols*[#], Michael Hopkins*

*f.lang@sussex.ac.uk

SPRU (Science and Policy Research Unit), University of Sussex, Brighton, UK

[#] Ingenio (CSIC-UPV), Universitat Politècnica de València, València, Spain

ABSTRACT

This paper proposes a novel approach to determine changes that occur as a result of collaborations that is intended to support knowledge integration. The approach combines and applies indicators of proximity, diversity, coherence and has potential applications in the study and evaluation of research collaborations.

The scientometric literature has been exploring the topic of knowledge integration and interdisciplinarity for more than a decade (Bordons, 2004; Zitt, 2005; Rafols, 2014). The paper builds on the line of research that seeks to develop measures of knowledge integration, namely diversity and coherence (Rafols, 2014). Successful exchange and integration of knowledge through collaboration not only requires disciplinary or cognitive diversity, as previously studied in the scientometric literature (Rafols & Meyer, 2009; Rafols, 2014) but also other dimensions linked to the social, cultural background of the individuals involved. Economic geographers have developed a framework, the proximity framework (Boschma, 2005), identifying five features that may be important for collaborative learning which are: cognitive, social, geographical, institutional, and organisational proximities.

The paper therefore proposes to use the diversity and coherence measures to not only look at diversity from a cognitive standpoint, but also apply it to the other proximities proposed in the Boschma framework. These indicators will capture the relationship occurring between individuals taking part in the research and the categories (proximity dimensions) that they are associated to. This paper reviews and integrates concepts from economic geography with the scientometric literature on interdisciplinarity to form a conceptual framework that the paper applies to an illustrative case study. In order to apply the framework, the paper develops indicators for diversity and coherence that can be applied to each of Boschma's five proximities. The illustrative case study looks at collaborations between individuals within a biomedical research project on Podoconiosis. The method aims to build not only indicators to look at diversity in collaboration, but also new ways of mapping relationships using the proximity framework.

INTRODUCTION

This paper proposes a novel approach to determine changes that result from collaborations aimed at supporting knowledge integration. The approach combines and applies indicators of proximity, diversity, and coherence. It has potential applications in the study and evaluation of research collaborations.

The science studies literature has been exploring the topic of knowledge integration and interdisciplinarity for decades, both in scientometrics (Bordons, 2004; Zitt, 2005; Rafols, 2014), research management (Grant & Baden-Fuller, 1995; Nooteboom, 2000) and geography (Maskell & Malmberg, 1999). Contributions have extensively discussed how different bodies

of knowledge are brought together within publication, individuals, university department or systems (Porter et al. 2007; Rafols et al. 2012 ; Zhou e al. 2012). The paper builds on the line of research that seeks to develop measures of knowledge integration, namely diversity and coherence (Rafols, 2014). Until now these measures of knowledge integration have mainly been used to look at disciplinary or cognitive differences.

The changes of successful exchange and integration of knowledge through collaboration is not only influenced by the disciplinary or cognitive diversity of the participants, as previously studied in the scientometric literature (Rafols et al., 2012; Rafols, 2014) but also other dimensions linked to the social, cultural background of those involved. A powerful framework has emerged from economic geography - the proximity framework (Boschma, 2005), that identifies five features that may be important for collaborative learning and hence collaborative knowledge integration which are: cognitive, social, geographical, institutional, and organisational proximities.

This paper therefore proposes to use diversity and coherence measures to not only look at diversity from a cognitive standpoint, but also apply it to the other proximities proposed in the Boschma framework. These indicators will capture the relationship occurring between individuals taking part in the research and the categories (proximity dimensions) that they are associated to.

This paper reviews and integrates concepts from economic geography with the scientometric literature on interdisciplinarity to form a conceptual framework that the paper applies to an illustrative case study. In order to apply the framework, the paper develops indicators for diversity and coherence that can be applied to each of Boschma's five proximities. The paper proposes to use diversity and coherence measures with the five dimensions, and it also uses the elements behind the diversity and coherence measures to create visualisations that has the potential to represent some complexity hidden behind the metrics. The illustrative case study looks at collaborations between individuals within a biomedical research project on Podoconiosis. The method aims to build not only indicators to look at diversity in collaboration, but also new ways of mapping relationships using the proximity framework.

2. ANALYTICAL FRAMEWORK

This section explores the framework developed by Rafols (Rafols & Meyer, 2010; Rafols, Porter, & Leydesdorff, 2010; Rafols, 2014) to study diversity, and shows how this can be associated to the theoretical framework developed by Boschma (2005) on the 5 proximities dimensions proposed. More specifically we propose to build a set of indicators, inspired by the literature on diversity and coherence, but applicable by researchers using the proximity dimensions. It will offer indicators that are descriptive of the system individuals are participating in; the categories they are associated to (described by the 5 proximity dimensions); as well as the bridges/flows through a dynamic indicator of coherence, that capture the relationship occurring between individuals taking part in the research and the categories that they are associated to.

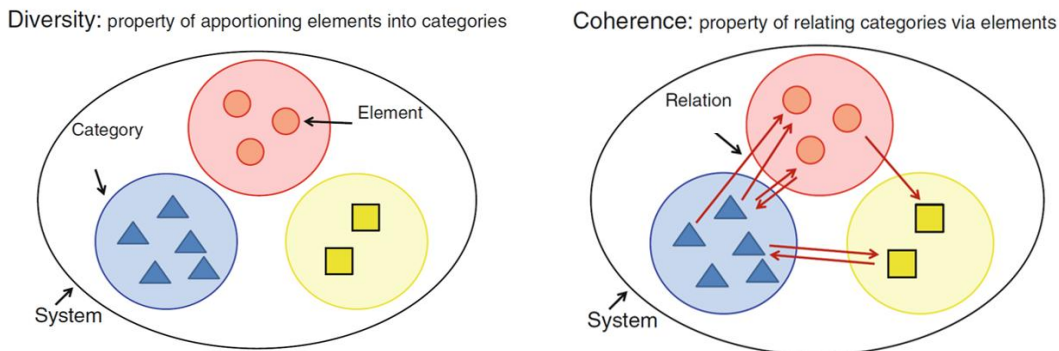
The contribution of the paper lies in the application of quantitative approaches to assess the diversity and coherence not only for cognitive aspects, but also to other social and organisational aspects in order to represent a wider range of aspects of research collaborations. The paper will first discuss how measures of diversity and coherence can be

operationalised for application to the study of collaborative research using the proximities dimensions, and then proposes an operationalisation for each of the proximity.

THE GENERAL MEASURES

The measures of diversity and coherence are based upon the classification of *elements* into *categories* as part of a specific *system* (Rafols, 2014). In this paper we are using *individuals* as *elements* of a formal research collaboration (i.e. a team funded to undertake collaborative project supported by a research grant) which forms the overall *system* being studied. For each of the five proximities, individuals (elements) will be assigned to different positions (categories), and the distance between individuals is characterised by the gap or difference between categories. **Figure 1** shows a representation of both diversity and coherence (Source: Rafols 2014).

Figure 1: Illustration of the definitions of diversity (left) and



For example, consider two collaborators working together on a common paper, one from the University of Sussex and the other from the Universitat Politècnica de València. In geographical terms they will be assigned to different categories, one to Brighton and the other in València (our two positions/ categories defined by the towns they work in) and the distance between these positions can be defined by the travelling time required to meet each other (time is chosen in preference to distance, which does not take into account the transportation infrastructures that support collaboration). The distance is approximately 6 hours by combination of air and public transport.

The diversity measure describes three inter-related properties: the **variety** or number of categories, the **balance** of elements across categories and the **disparity** among categories. For example, following the geographic example above, a collaboration between two authors in the same city with another in a second city, is characterised as having less variety and balance than three authors in three cities, while the latter case has more disparity than the former.

Rao-Stirling diversity

$$\sum_{i,j(i \neq j)} p_i p_j d_{ij}$$

In addition to the positions of the elements, another property we are interested in is the relationship between actors. In the case of interactions, we need to take into account both the number of interactions and their intensity. This is captured by the coherence measure. The

coherence measure describes three interrelated properties: the **density** of interactions, the relative **intensity** of the relations and the **disparity** across categories bridged by relations.

Coherence (□□□□□□□□)

$$\sum_{i,j(i \neq j)} i_{ij} d_{ij}$$

Thus these two measures used along the five dimensions in the proximity framework capture the **position** of each element/individual within the system studied, but also the **bridges/flows** that emerged between the different dimensions. Later in the paper we will look at how we can apply these measures to study an illustrative case study of collaboration within a biomedical research project. For each indicator we can calculate position and flows before and after the research project started in order to capture the impact of the project on both the positions of the actors and the flows between actors. The paper follows on to describe further the operationalisation of the proximities in 1) the position and 2) the flows indicators.

The attribution of a specific position to an element (or individual) is made for each of the proximity. The indicator shows the distribution of the elements under study within specific sets of categories. The indicator is static as it provides a picture at a given time of the distribution of individuals across categories. To display the position of actors within categories we use an indicator based on the Rao-Stirling measure of diversity, which includes a component about the distance between individuals and the proportion of individuals within the categories in the system.

There are five categories, one for each dimension of Boschma's proximity framework: The social, the cognitive, the organisational, the institutional and the geographical. The five categories are here operationalised in five distinct attributes of distance. These require distinct strategies for operationalisation.

For three proximities (geographical, organisational, and institutional) the association of elements to each category is defined depending on where the individuals (elements) work (i.e. the geographical location of the workplace, the organisation employing the person and the type of institution) (Ponds et al. 2007, Hardeman et al. 2015). Another category is defined with reference to the knowledge base on which individuals build within the collaboration (the cognitive dimension) and is captured through the citations from scientific publication for each individual (Porter et al. 2007, Rafols 2014). The final proximity (the social dimension), assumes that each individual is both a distinct element and a different category. While we have described how the proposed operationalisation assigns elements to categories for the analysis, the diversity indicator also includes a distance component to represent disparity between each categories.

3. OPERATIONALISATION OF MEASURES OF DIVERSITY AND COHERENCE

In order to operationalise the measures, we need to compute the proportion of elements into categories (p_i) and define the proximities or distances between categories. This has to be done for each of the five analytical dimensions.

Social distance

For the **social distance**, it is proposed to take into account whether individuals know each other before the project started. Two individuals will be considered close at the start of the

project depending on two factors - whether they knew each other before the project (they would be considered quite close) and whether they worked together previously (in which case they will be considered very close - if they had worked together before and were prepared to work together again in a new project). We describe the distance between locations in terms of a categorical variable:

- Do not know (4/4)
- Know a bit (3/4)
- Have worked together before (0/4)

Cognitive distance

The **cognitive distance** is based on the journals cited in papers individuals considered authored. Each individual is associated with a number of journals they cite, and the distance between two individuals will be based upon the journals they cite and whether these journals are considered similar or not. The similarity between journals is defined by a similarity matrix, based on citation patterns for individual journals. This similarity matrix is produced on the basis of citations between the web of science indexed journal (for the last 5 year period), and has been kindly provided to us by the OST (Paris). This data is used to calculate cosines similarity between each pair of journals which is used as the similarity matrix between journals.

This journal similarity matrix is used together with citations patterns of individuals to calculate distances between individuals taking part in the project. This is performed using a method proposed by Zhou et al. (2012) who describe a way to compute a similarity-weighted cosine measure. The similarity-weighted cosine measure (which is in a normalised form) is defined as follow:

$$\varphi(X, Y) = \frac{\varphi_{X,Y}}{\sqrt{\varphi_X \varphi_Y}} = \frac{\sum_{i,j=1}^N p_{X_i} p_{Y_j} S_{ij}}{\sqrt{\left(\sum_{i,j=1}^N p_{X_i} p_{X_j} S_{ij}\right) \cdot \left(\sum_{i,j=1}^N p_{Y_i} p_{Y_j} S_{ij}\right)}}$$

This measure enables us to provide a similarity measure between two individuals depending on their cognitive background, which cited journals as a proxy.

Geographic distance

For the **geographical distance**, individuals are assigned to a geographic location. The geographic location is assigned depending on the time spent by a person at a specific location. In some cases, this can be different from their affiliation (based on data reported by individuals, for example at interview). Thus the geographical and organisational distances can be based on different data, while organisations' addresses are used to calculate geographical distances. As previously noted we use travelling time as a proxy of geographical distance between two individuals working at different given locations.

We describe the distance between locations in terms of a categorical variable:

- Same department (3 minutes) (0/5)
- Same university, same campus (up to 15 minutes' walk) (1/5)
- Same city/metropolis (up to 2 hours) (2/5)

- Same region/country (up to 4-5 hours by train) (3/5)
- Same continent (flight or long train needed) (4/5)
- Other continent (5/5)

Organisational distance

For the **organisational distance**, individuals are assigned to an organisation, the organisation they work in. There are different levels of organisational integration which we take into account when defining the distance (whether the individuals work in the same department or centre, or if an individual has a visiting status in an organisation).

We describe the distance between the organisations in terms of a categorical variable:

- Same department or centre (0/2)
- Same organisation (1/2)
- Different organisation (2/2)

If the person has a **visiting status** in an organisation he/she will have a (-1/2) to correct for the status, as these individuals may be considered closer in organisational terms than people who are completely external to the organisation.

Institutional distance

Finally, for the **institutional distance**, we use previous literature in order to define distances between given institutions. As our example focuses on a biomedical research projects we consider six different type of institutions which has been previously identified in the literature (Rotolo et al., 2015): those involved in higher education/ research (e.g. universities), hospitals, governmental organisations, non-governmental organisations, industry, and university hospitals. In this latter category we differentiate between individuals mainly working as clinicians in university hospitals (referred to as working in Hosp/Univ) and those mainly working as researchers (referred to as working in Univ/Hosp) given the different requirements attached to these roles. In order to identify distances between institutions, we consider the overlap over the general missions between these institutions. We consider whether these institutions' main objective are oriented towards commercialisation, Care, Open science, Education and Policy (Llopis & D'Este, 2016). The following table whether each institution has one or more of the following mission, with a yes or no answer represented by a binary attribute.

Table 1: Overlap of missions between different institutions

	Res&Edu	Hosp	GO	NGO	Industry	Univ/Hosp	Hosp/Univ
Commerc.	0	0	0	0	1	0	0
Care	0	1	0	0	0	1	1
Open Sc	1	1	1	1	0	1	1
Education	1	0	0	0	0	1	1
Policy	0	0	1	1	0	0	0

We use each columns as a vector of binary attributes (see table 1). The above table can then be interpreted as a contingency table for binary attributes. Using the symmetric binary dissimilarity method (Han, Kamber, & Pei, 2012, pp. 70–71) we can compute the table below :

Table 2: Institutional distance defined between pairs of institutions (1)

	Res & Edu	Hosp	GO	NGO	Industry	Univ/Hosp	Hosp/Univ
Res & Edu	0	0.4	0.4	0.4	0.6	0.2	0.2
Hosp	0.4	0	0.4	0.4	0.6	0.2	0.2
GO	0.4	0.4	0	0	0.6	0.6	0.6
NGO	0.4	0.4	0	0	0.6	0.6	0.6
Industry	0.6	0.6	0.6	0.6	0	0.8	0.8
Univ/Hosp	0.2	0.2	0.6	0.6	0.8	0	0
Hosp/Univ	0.2	0.2	0.6	0.6	0.8	0	0

The distance ranges from 0 to 0.8 between pairs of organisation. There are a few concerns with this similarity matrix as we would like to consider each institutions to be different from one another, thus GO and NGO must be superior to 0, and people working on the research side (Univ/Hosp) from the university-hospitals must be differentiated from those working as clinicians (Hosp/Univ). Also we would like to readjust the measures between Universities, Hosp, Univ/Hosp and Hosp/univ. As the primary focus of University-Hospitals and universities is teaching and open science, they should be closer than the ones that have their main focus on care (Hospitals and clinicians at university hospitals). Univ/Hosp and Hosp/Univ are different because the first is slightly more focused on open science and the latter is primarily focused on care. Thus the distance measure will be slightly modified to take into account this aspect.

FLAWS/BRIDGES (USING THE COHERENCE INDICATOR)

In addition to exploring ways to assess diversity, we also provide an operationalisation for assessing coherence that looks at the flows occurring within the project between the different categories across each of the five proximities. The bridges or flows are being represented through the coherence indicator that include both factors for **distance** (which uses the distance measures discussed in the previous section) and **intensity** as introduced below.

The intensity of the flows are based on indicators of personal interactions made by individuals. The intensity measure is therefore defined by the frequency of interactions (i.e. whether these are daily, weekly, monthly, bi-annually, or yearly interactions). The measure of intensity is different to the social proximity set out above because here we are concerned not with how acquainted individuals are but by the frequency of interaction which is used as a proxy for intensity of collaboration.

Intensity measures

The scale of intensity can be derived from the frequency of the interaction, this is a measure of personal interaction . For example:

- no meeting (0)
- yearly meeting (1/5)
- every 6 month meeting (2/5)
- monthly meeting (3/5)
- weekly meeting (4/5)
- daily meeting (5/5)

4. APPLICATION TO A RESEARCH PROJECT ON PODOCONIOSIS

The indicators presented above were purposefully presented in a general manner in order to introduce a novel way to study collaboration. The last part of the paper aims at applying the developed indicators to a specific case. The case follows collaboration within a funded research project, in this case we will focus on a research project aiming at developing the understanding of a specific neglected disease, podoconiosis. Podoconiosis is a relatively under studied non-infectious neglected tropical disease which is characterised by the swelling of feet or lower part of the leg in affected individuals (Deribe, Tomczyk, & Tekola-Ayele, 2013). It is associated with social stigma and is also causes significant problems by reducing the economic activity of sufferers. The focal research project resulted in a substantial boost to the number of publications on this topic as well as increasing substantially the number of researchers working in this field.

The data relies on both publication and interview data gathered among individuals participating in the research project. Publication data were retrieved through the Web of Science and are mainly used to generate indicators and maps of cognitive proximity, as already explored in previous literature (Rafols, 2014). Interview data consists of gathering data about the other proximities such as organisations, institutions, geographical location (which can be crossed checked with the publication data), social relationships, but also data about intensity of interactions.

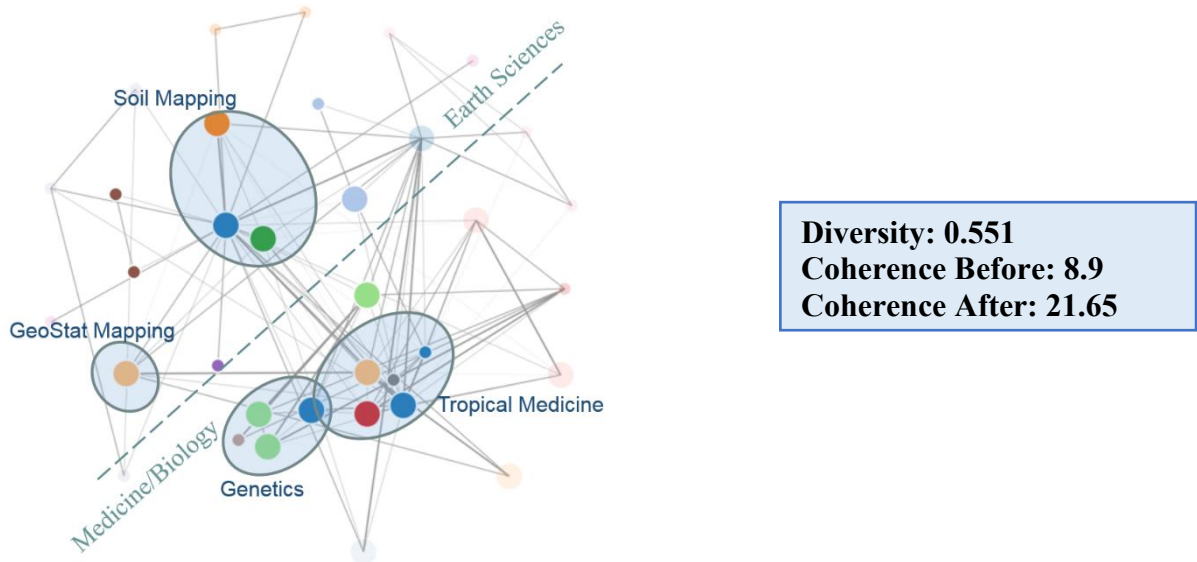
VISUALISATION

Diversity and coherence include different elements, such as distance between categories (included in both measures), the intensity of links for the coherence measures and the proportion of elements in each category (included in the diversity measure). For each of the measures, namely diversity and coherence, two of these aspects are integrated into the measure and therefore the analyst loose part of the complexity of the information held into the single metric. Thus the analysis of the proximity dimensions can be performed using both the measures (i.e. coherence and diversity) explained above, together with the visualisation that enables the user to have a better understanding of the single metric. For instance, as discussed above, the diversity is based upon how elements are distributed into categories, and have attributes such as variety, balance, and disparity. In the same way coherence has properties such as density of interactions, intensity of relationships, and disparity across categories as well as how they are bridged.

The visualisations are represented in a two dimensional space and show distances between individual elements and links (intensity is displayed by the thickness of the line) between these individuals. The distribution of elements enables the analyst to identify categories. This can be done by using both the information given in interviews (for social and organisational proximities) and information held in raw data (for the cognitive side). Figure 2 to 7 shows such representations based on data collected in the Podoconiosis project case study, for each of the proximity dimensions using part of the metrics introduced above.

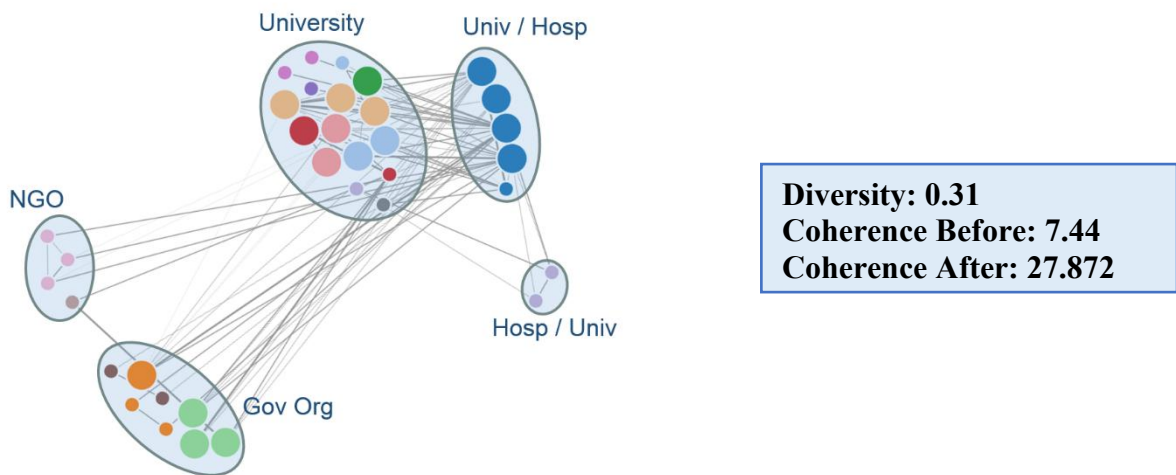
COGNITIVE PROXIMITY

Figure 2: Links between individuals with cognitive node positioning



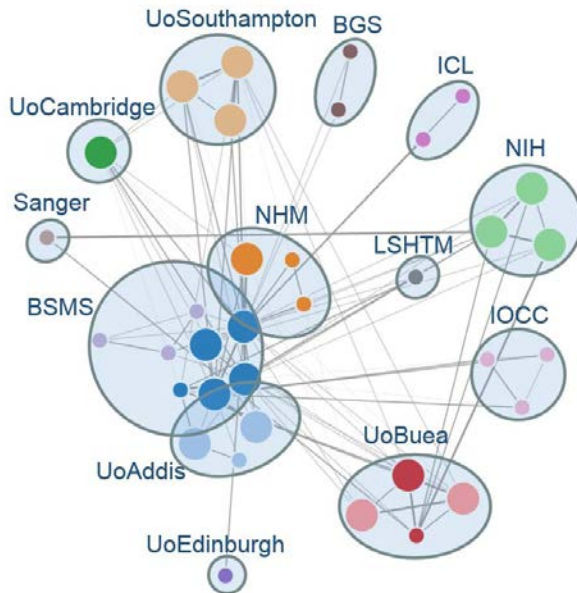
INSTITUTIONAL PROXIMITY

Figure 3: Links between individuals with institutional node positioning



ORGANISATIONAL PROXIMITY

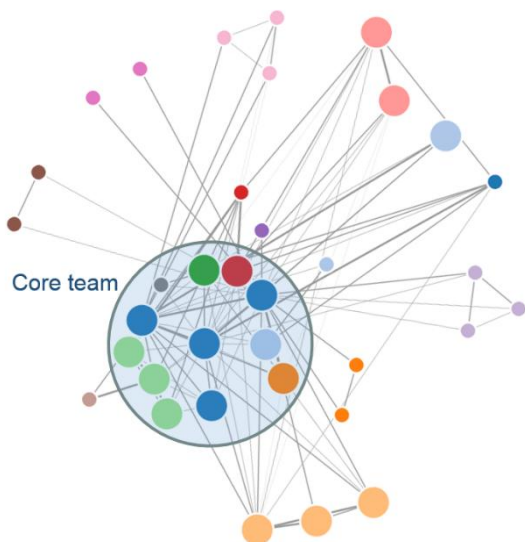
Figure 4: Links between individuals with organisational node positioning



Diversity: 0.905
Coherence Before: 20.6
Coherence After: 71

SOCIAL PROXIMITY

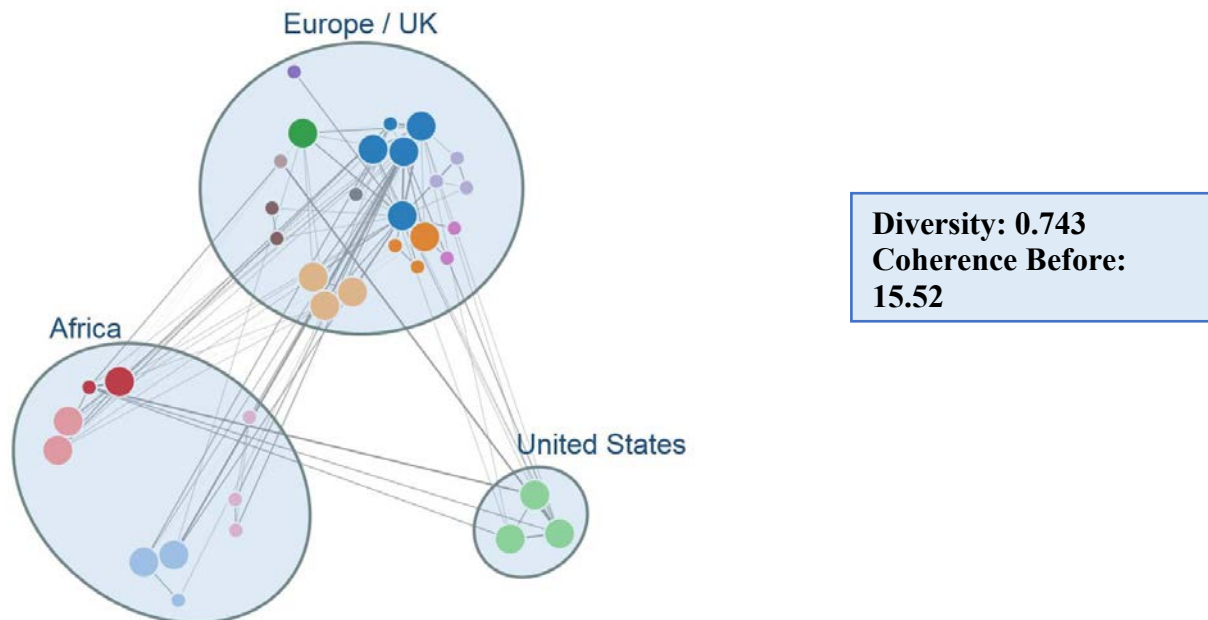
Figure 5: Links between individuals with node positioning



Diversity: 0.924
Coherence Before:
14

GEOGRAPHIC PROXIMITY

Figure 6: Links between individuals with geographic node positioning

**5. CONCLUSIONS**

The paper concludes with a discussion about the suitability of the proposed tool to assess potential knowledge integration through collaboration, and its strengths and limitations.

REFERENCES

- Bordons, M., Morillo, F., & Gomez, I. (2004). Analysis of cross-disciplinary research through bibliometric tools. In: Moed, Glanzel, & Schmoch (Eds.), *Handbook of quantitative science and technology research* (pp. 437–456). Dordrecht: Kluwer.
- Boschma, R. (2005). Proximity and Innovation: A Critical Assessment. *Regional Studies*, 39(1), 61–74. doi:10.1080/0034340052000320887
- Grant, R. M., & Baden-Fuller, C. (1995). A KNOWLEDGE-BASED THEORY OF INTER-FIRM COLLABORATION. *Academy of Management Proceedings*, 1995(1), 17–21. doi:10.5465/AMBPP.1995.17536229
- Han, J., Kamber, M., & Pei, J. (Computer scientist). (2012). *Data mining: concepts and techniques*. Elsevier/Morgan Kaufmann.
- Llopis, O., & D'Este, P. (2016). Beneficiary contact and innovation: The relation between contact with patients and medical innovation under different institutional logics. *Research Policy*, 45(8), 1512–1523. doi:10.1016/j.respol.2016.03.004
- Maskell, P., & Malmberg, A. (1999). Localised learning and industrial competitiveness. *Cambridge Journal of Economics*, 23(2), 167–185. doi:10.1093/cje/23.2.167
- Nooteboom, B. (2000). Learning and innovation in organizations and economies.
- Rafols, I. (2014). Measuring SKnowledge Integration and Diffusion: Measures and Mapping of Diversity and Coherencecholarly Impact. In Y. Ding, R. Rousseau, & D. Wolfram (Eds.), *Measuring Scholarly Impact* (pp. 169–190). Cham: Springer International Publishing. doi:10.1007/978-3-319-10377-8
- Rafols, I., & Meyer, M. (2010). Diversity and network coherence as indicators of

interdisciplinarity: Case studies in bionanoscience. *Scientometrics*, 82(2), 263–287. doi:10.1007/s11192-009-0041-y

Rafols, I., Porter, A. L., & Leydesdorff, L. (2010). Science overlay maps: A new tool for research policy and library management. *Journal of the American Society for Information Science and Technology*, 61(9), 1871–1887. doi:10.1002/asi.21368



Using a network-based approach to identify interactions structure for innovation in a low-technology intensive sector¹

Camille Aouinait^{*}, Benedetto Lepori^{**}, Danilo Christen^{***}, Christoph Carlen^{***} and Dominique Foray^{*}

^{*}camille.aouinait@epfl.ch; dominique.foray@epfl.ch
College of Management of Technology (CEMI), Ecole Polytechnique Fédérale de Lausanne, Odyssea, 1015 Lausanne (Switzerland)

^{**} benedetto.lepori@usi.ch
Research service and Centre for Organizational research, University of Lugano, Via Lambertenghi 10a, 6904 Lugano (Switzerland)

^{***} danilo.christen@agroscope.admin.ch; christoph.carlen@agroscope.admin.ch
Research group in fruits crops in alpine region, Agroscope, Route des Eterpys 18, 1964 Conthey (Switzerland)

ABSTRACT

Knowledge transfer in the agricultural network is realized through interactions between stakeholders, inducing innovation development and diffusion. The aim of the paper was to trace interactions in the Swiss apricot sector. Identification of collaborations using face-to-face interviews of knowledge producers and knowledge users were conducted. The study showed that informal collaborations are exclusively used to transfer knowledge and create innovation. Personal ties have been established between internal actors of the value chain (e.g. professionals like producers, transformers and wholesalers). External partners like public research organizations have created strong ties with agricultural stakeholders. However, the spatial proximity does not guarantee higher rate of collaborations. The links with the Universities of Applied Sciences, closely located, are sparse. Hence, in order to warrant innovation success, spatial proximity has to be balanced with organizational proximity. Despite the educational background of producers, there are a few connections with universities. Human capital formation and education in the agricultural sector should be examined to design innovation policy. Besides, the public research center for agriculture catalyzes knowledge transfer and facilitates innovation adoption. A suitable ecology of actors through the value chain from research to application is necessary. Furthermore, productive interactions should be investigated to identify the efficiency of knowledge and innovation transfer mechanisms and potential gaps in this process.

BACKGROUND

In the context of small enterprises, implementation of innovation is required to gain competitive advantages and therefore, given the limited R&D capabilities of companies, interactions with public research institutes are crucial (Cohen, Nelson, & Walsh, 2002). In the literature, formal and informal interactions between stakeholders are commonly opposed in terms of their importance to drive knowledge and innovation (Agrawal & Henderson, 2002; D'Este & Patel, 2005; Lissoni, 2001; Reagans & McEvily, 2003; Roper, Vahter, & Love, 2013).

¹ This work was supported by the European Commission within the TRAFON project (FP7-KBBE project number 110874)

In 2014, 95.5% of Swiss apricots are produced in one region, the canton of Valais. 90 professionals and 300 traditional farmers produced 7,700 tons of 70 varieties spread on 676.5 hectares (OFAG, 2015; Roher, 2012; Valais-Wallis.Promotion, 2015).

According to the taxonomy of sectors by Pavitt (1984) agriculture is a supplier-dominated sector. Consequently, innovation mostly comes from suppliers of inputs (e.g. fertilizers, plants, pesticides) and from agricultural research. Basic and applied research is achieved in different firms or public organizations such as the Agroscope, Swiss center of excellence for agricultural research, Research Institute of Organic Agriculture (FiBL) and academic sites such as federal schools, cantonal universities and Universities of Applied Sciences (UAS).

PURPOSE

The goal of the paper is to investigate the characteristics of the interaction process between firms and public research related to agricultural production in a peripheral region of Switzerland. An analysis of the collaborations occurring in the network is conducted. The focus is on three aspects: (i) the importance of informal interaction channels used by apricot chain actors for innovation, (ii) the role of geographical proximity and (iii) the presence of knowledge transfer catalysts. Then, types of innovation implemented are examined.

SWISS APRICOT NETWORK ANALYSIS

In the apricot network, information flows convey between nodes represented by stakeholders that are internal to the value chain (e.g. professionals like nurseries, producers, transformers, retailers) and external like research units in universities or public organizations, professional associations and public bodies.

FORMAL AND INFORMAL KNOWLEDGE TRANSFERS

Literature on collaborations between university and industry focused on formalized interactions using codified transfers (e.g. patents, co-publications, licenses) (Foray & Lissoni, 2010; Rossi & Rosli, 2013). Nevertheless, several studies demonstrated the low importance of these channels to transfer knowledge (e.g. 10% via patents) especially in applied research (Agrawal & Henderson, 2002; Ponomariov & Craig Boardman, 2008). Using informal relationships for transferring assets and producing impacts on firms performance is substantial (D'Este & Patel, 2005). Moreover, innovations in agriculture are hardly to observe and to measure. Informal interactions should bring information on knowledge transfer structure. Thus, the first hypothesis to test is the following;

H1: Informal interactions have a higher impact than formalized interactions on regional innovation development.

GEOGRAPHICAL PROXIMITY

Geographical proximity between actors can improve innovation efficiency and knowledge transfer (Rallet & Torre, 1998). The Swiss apricot production is regionally concentrated; the actors are close to each other. One of the initial mandates of UAS creation was the accomplishment of applied research activities and the support of local economy (Jongbloed, 2010; Lepori, Huisman, & Seeber, 2012). Hence, the second hypothesis assumed is:

H2: Short distance in terms of geographical proximity induces more collaboration between apricot actors and research centers in universities and public institutions than bigger distance.

ORGANIZATIONAL PROXIMITY

Rallet and Torre (1998) defined organizational proximity as a “collective and technical culture” that leads to homogenization of individuals’ behavior in a firm. It can be seen as close missions or mandates undertaken by institutions. Organizational proximity overcomes geographical proximity and is stronger for knowledge transfer, technology transfer and innovation diffusion (Lissoni, 2001).

Besides, Klerkx and Leeuwis (2009) stated that there are catalysts that enhance innovation diffusion. Collaboration between private sector and research organizations improves research efficiency and its impact on society, especially in the sector of traditional food production. It drives us to the third hypothesis following;

H3: There are more collaboration between apricot actors and research actors presenting high organizational proximity.

TYPE OF IMPLEMENTED INNOVATIONS

Apricot supply chain actors create and adopt different types of innovations through collaborations. The study identifies the innovations implemented in the last three years.

METHODS AND MATERIALS

Data for the study were partially derived from the EU-FP7 project on Traditional Food Network to improve the transfer of knowledge for innovation (TRAF00N). The project focuses on traditional food producing by Small and Medium-sized Enterprises (SMEs) of apricot and berries actors in Switzerland. Data on producers, transformers, retailers, and other relevant stakeholders of the sectors were identified. For the current study purpose, an investigation of the knowledge diffusion arising in the low-technology sector of apricot production was conducted through existing collaborations. Interviews were performed to draw the network. The list of respondents was chosen from the TRAF00N project’s data.

APRICOT SUPPLY CHAIN DESCRIPTION FOR THE NETWORK ANALYSIS

Interactions occurring in the network follow iterative loops, allowing constant improving of goods, management system, services or marketing strategy. Producers interact with national and foreign research institutes for different purposes such as the follow-up of new products (e.g. varieties, phytosanitary products, irrigation system) and advice about crop management. They acquire new varieties from nurseries and plant breeders. After the harvesting activities, the fruits are either sold by the farmers (in fresh and/or processed state) or traded to wholesalers and transformers. Products are distributed to consumers by different channels: mass distribution, retail stores, specialized shops (fruits and vegetables shops, organic shops, distilleries) and catering. The sector has been renewed since 20 years to meet the whole chain expectations; longer harvesting period, diversified taste preferences by consumers, fast maturity inducing transport issues.

19 interviews have been conducted (52.7% of the sample targeted), including 40% of producers, 20% of transformers, 14% of traders and 10% of wholesalers amongst others. The less important distribution channel used is direct selling. 26% do not use it, 37% rarely, 21% occasionally and 16% frequently. The mean apricot surface of SMEs interviewed is 15.82 hectares per farm (variation between 4 to 54 hectares). There are mainly incumbent companies established in average in 1962. Management of labor force is highly variable

according to the season. The average number of employees in winter is 13.83 full-time equivalent and 40.06 full-time equivalent in the summer period.

HYPOTHESES TEST

The work uses the methodological approach of the FP7 European project Social Impact Assessment Methods through Productive Interactions (SIAMPI) that was developed to assess the impact of research using the connectivity between supply chain actors (e.g. direct, indirect and financial interactions).

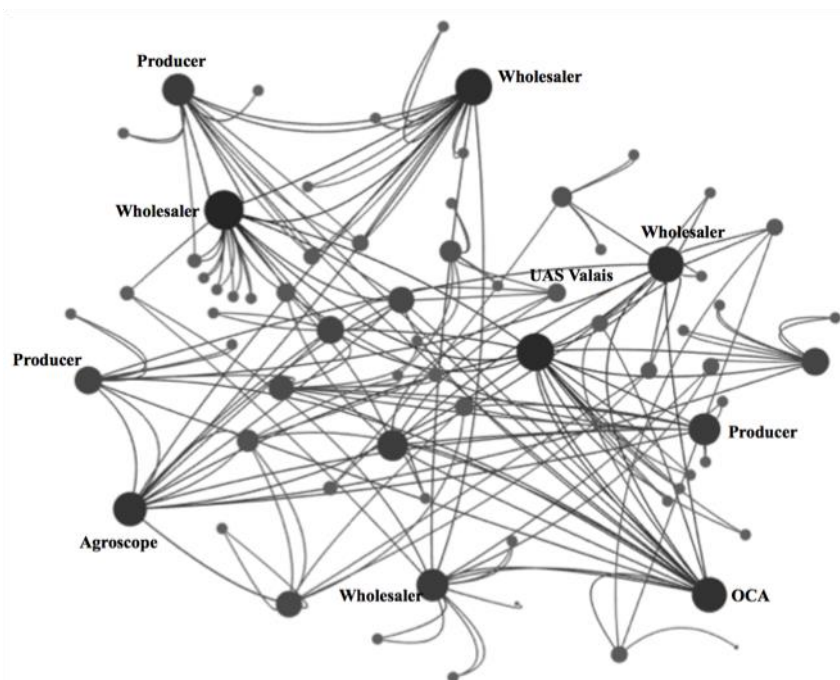
Semi-structured surveys were used during the interviews. Collaborations about innovation were asked to be specified with the name of collaborator, collaboration type, collaboration way (e.g. emails exchanged, phone, visits), reciprocity (bilateral versus multilateral) and resources exchanges (information, material) (Burt, 1997; Nicolaou & Birley, 2003).

FINDINGS

SWISS APRICOT NETWORK ANALYSIS

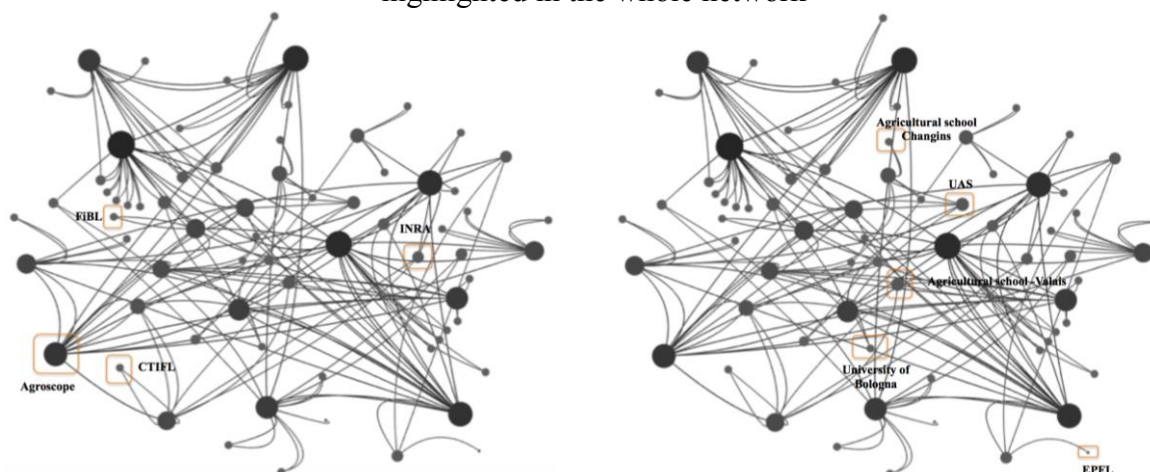
Connectivity between actors is presented in figure 1. There are 70 nodes corresponding to actors and 193 edges representing the collaborations cited by the respondents.

Figure 1: Collaborations in the Swiss apricot network



Labels of important nodes are reported in figure 1. The bigger the nodes, the bigger the degree centrality, meaning high number of connections with neighbors. Wholesalers present the highest degree centrality: 23 (VS Fruits), 21 (Fruitex), 20 (Alpfruits) and 19 (Pitteloud Fruits). The Cantonal Office of Arboriculture (OCA) and the public research organization (Agroscope) have respectively a degree of 18 and 17. A lot of respondents named these actors as collaborators. Position of research organizations and universities in the network can be visualized in figure 2.

Figure 2: Research institutes (left graph) and schools and universities (right graph) highlighted in the whole network



The edges going in and out of the research actors in public organization and universities are scattered. Nonetheless, Agroscope, located at the down left corner of the figure 2, dominates researchers ties with fruit actors.

Collaborations cited by stakeholders regarding business interactions and transfer of innovations are presented in table 1.

Table 1. Swiss apricot network interactions on knowledge and innovation transfer

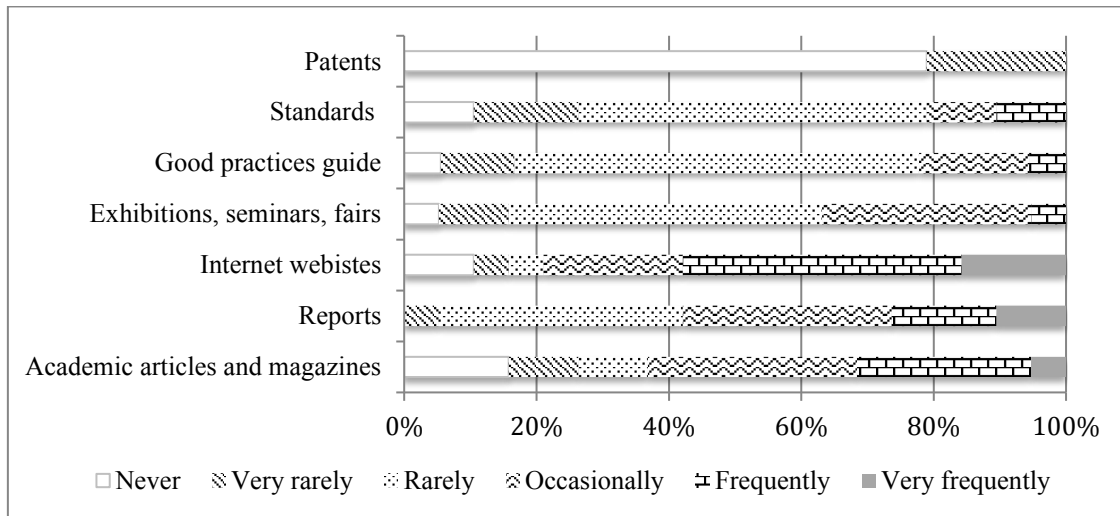
Categories of actors	Number of citations	Percentage
Nurseries	13	7.8
Producers	84	50.6
Transformers	8	4.8
Wholesalers	60	36.1
Customers (retail and specialized stores)	1	0.6
<i>Total</i>	<i>166</i>	<i>100%</i>

Producers are strongly connected, with 84 ties mentioned. This is the main category cited (50.6%), followed by wholesalers (36.1%). The actors exchange information, materials like fruits and packaging. Innovations diffuse among the network.

FORMAL AND INFORMAL TRANSFERS

Communication tools used to exchange information and materials are mostly emails and phone. Face-to-face communication and field visits increase during summer. The figure 3 reports the written sources of information that stakeholders consult to acquire knowledge about innovation. Patents are almost never consulted. Standard and good practices guide are rarely consulted (52.6% and 61.1% respectively). Fairs and exhibitions, rated “rarely” (47.4%), represent the opportunity to create direct contact with visitors (professionals and consumers), to diffuse innovation and to gather information about other participants’ innovations. The other sources are the most frequent used ones, especially Internet websites (42.1% frequently used). Reports and academic articles are read occasionally (once every two month) because of revues subscriptions (31.6% each).

Figure 3: Frequency of sources of written information used to learn more about innovation



According to these findings, informal interactions are the most exploited channels to transmit information, materials and other assets conducting to innovation. Hypothesis 1 is confirmed; informal interactions have a higher impact than formalized interactions on regional innovation.

GEOGRAPHICAL AND ORGANIZATIONAL PROXIMITY

Geographical proximity and number of collaborations between apricot actors and research actors are presented in table 2.

Table 2. Swiss apricot network interactions on knowledge and innovation creation

Categories of actors	Geographical proximity ²	Number of citations	Percentage
Research and technical institutes		15	44.1
<i>Agroscope</i>	+	10	29.4
<i>FiBL</i>	-	2	5.9
<i>INRA</i>	-	2	5.9
<i>CTIFL</i>	-	1	2.9
Policy and professional associations		11	32.4
<i>OCA</i>	+	10	29.4
<i>IFELV</i>	+	1	2.9
Universities		8	23.5
<i>EPFL</i>	-	1	2.9
<i>UAS Valais-Wallis</i>	+	3	8.8
<i>University of Bologna</i>	-	1	2.9
<i>Agricultural school of Valais</i>	+	2	5.9
<i>Agricultural school Changins</i>	-	1	2.9
<i>Total</i>		34	100%

² Geographical proximity is defined as close (+) if the organization or firm is located in the Valais region, far (-) otherwise

There are main links established with research and technical institutes, cited at 44.1% (15 citations), policy body and associations (11 citations).

The proximity column demonstrates that a short distance between professionals of the apricot sector and external actors like research institutes, associations and universities does not guarantee more collaboration. Indeed, Agroscope, OCA, IFELV, UAS Valais-Wallis and agricultural school of Valais are close (noted “+” in table 1). However, UAS has few connections with stakeholders (3 citations). Agroscope predominates collaboration activities with 29.4% of the total collaborations reported. Therefore, hypothesis 2 is not totally confirmed. Geographical proximity is not sufficient to establish efficient interactions between stakeholders. This result is surprising as 28% of the respondents are training agronomist and oenologist engineers that studied in the UAS. This can probably be exploited by the specific mandates of the different institutions.

The high implication of Agroscope and OCA in collaborations reveals a catalyst role. The institutions transfer the knowledge by translating research outputs into ready-to-use information and innovation to the practice. They have closer organizational proximity to the respondents than other institutes. Thus, hypothesis 3 is confirmed.

TYPES OF IMPLEMENTED INNOVATIONS

Firms have been asked to specify the innovations implemented in the last three years. 16 firms have implemented at least one product innovation, mostly new varieties. It was developed in collaboration with public research institute. Seven firms implemented process innovations (e.g. juices, dried fruits, liquors). Six firms applied organizational innovations and nine firms are concerned by marketing innovations. A firm reported trends to “go close to the consumers” and “fresh and local products”. To “ensure economic valorization”, e-shop platform should be improved. Explanations about varieties like hedonic criteria and maturity date would be provided. Lastly, distillers are limited in innovation possibilities because of legal requirements about advertising. Hence, one firm launched a social-product innovation to meet evolution of consumption trends.

DISCUSSION

NETWORK ANALYSIS

To measure the impact on innovation, data on market value and positioning of the sector should be collected. Further research should identify which productive interactions drive innovation. Recipients’ behavior may change depending on the interactions’ efficiency. Besides, strong and weak ties have both the potential in contributing to innovate at the sector level. On one hand, exploring the complementarities of Agroscope and UAS should be realized. It would facilitate implementation of policy measures in order to foster innovation development. On the other hand, secondary ties used for knowledge flows should be analyzed in order to collect information on the network structure and consequently innovation capacity of the sector.

In their study about university and industry relationships, Bercovitz and Feldman (2006) assumed that informal links potentially precede formal links, especially in applied research. Thus, formalizing and promoting current informal interactions may be the next step to enhance network cohesion. Finally, human capital formation and education has been hardly studied and should be examined in the setting of agricultural sector.

INFORMAL COLLABORATIONS AND PROXIMITY

In the Swiss apricot network, innovation disclosure is realized through informal interactions. These findings are in line with the work by Poncet et al. (2010) who showed that formal intermediaries do not prevail the innovation networks in irrigated crops like vegetables, sugar cane, sugar beet and maize.

Moreover, the study highlighted the importance of combining organizational proximity and geographical proximity in order to efficiently promote network ties. A dedicated public research organization catalyzes these links (Agroscope). The results showed a high rate of collaborations between this institute and apricot professionals, but a few links with the University of Applied Sciences. In the sector of fresh products, activities conducted by Agroscope are closer to the producers' concerns and expectations than the UAS, which is closer to actors in processed products. Besides, the sample included 40% of producers. This bias has to be taken into account in the whole picture.

INNOVATION TYPES

The study identified product innovation as the major type of innovation adopted. Nevertheless, SMEs reported the willingness to allocate resources for marketing innovations. Strategically, small firms adopt product differentiation and market segmentation in a situation of information asymmetry on the markets and imperfect competition (Smith, 1995). Niche markets could be a solution for this small-scale sector. Hence, researchers are working in this area to meet expectations of producers and consumers, which ask for local apricots and targeting a broad set of taste preferences.

CONCLUSIONS

The research aims to bring insights for innovation and the policy identifies knowledge transfer processes conducting to innovation impacts. Consequently, a suitable ecology of actors through the value chain from research to application is necessary. Innovation success is warranted with implication of different actors.

This study showed that there are plentiful interactions within the Swiss apricot sector between professionals and public research organizations. Only informal interactions based on personal links are established. Universities of Applied Sciences are not the main partners of the sector. On the contrary, the national public research organization Agroscope display more collaboration projects with apricot professionals. Besides geographical proximity, organizational proximity is needed to boost network interactions.

REFERENCES

- Agrawal, A., & Henderson, R. (2002). Putting Patents in Context: Exploring Knowledge Transfer from MIT. *Management Science*, 48(1), 44–60.
- Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2002). Links and Impacts: The Influence of Public Research on Industrial R&D. *Management Science*, 48(1), 1–23.
- D'Este, P., & Patel, P. (2005). University - Industry linkages in the UK : What are the factors determining the variety of university researchers' interactions with industry? *DRUID Tenth Anniversary Summer Conference 2005 on Dynamics of Industry and Innovation: Organizations, Networks and Systems* (p. 29).

- Foray, D., & Lissoni, F. (2010). University Research and Public-Private Interaction. *Handbook of The Economics of Innovation* (Vol. 1, pp. 276–314).
- Jongbloed, B. (2010). Chapter 2: The regional relevance of research in universities of applied sciences. In S. Kyvik & B. Lepori (Eds.), *The Research Mission of Higher Education Institutions outside the University Sector* (pp. 25–44).
- Klerkx, L., & Leeuwis, C. (2009). Establishment and embedding of innovation brokers at different innovation system levels: Insights from the Dutch agricultural sector. *Technological Forecasting and Social Change*, 76(6), 849–860.
- Lepori, B., Huisman, J., & Seeber, M. (2012). Convergence and differentiation processes in Swiss higher education : an empirical analysis. *Studies in Higher Education*, 37–41.
- Lissoni, F. (2001). Knowledge codification and the geography of innovation: The case of Brescia mechanical cluster. *Research Policy*, 30(9), 1479–1500.
- OFAG. (2015). *Les cultures de fruits et de raisin de table de la Suisse en 2015*.
- Poncet, J., Kuper, M., & Chiche, J. (2010). Wandering off the paths of planned innovation : The role of formal and informal intermediaries in a large-scale irrigation scheme in Morocco. *Agricultural Systems*, 103(4), 171–179. Elsevier Ltd.
- Ponomariov, B., & Craig Boardman, P. (2008). The effect of informal industry contacts on the time university scientists allocate to collaborative research with industry. *Journal of Technology Transfer*, 33(3), 301–313.
- Rallet, A., & Torre, A. (1998). On Geography and Technology. The Case of Proximity Relations in Localised Innovations Networks. *Clusters and Regional Specialisation* (Steiner M.). Pion Publication, London.
- Reagans, R., & McEvily, B. (2003). Network Structure and Knowledge Transfer : The Effects of Cohesion and Range. *Administrative Science Quarterly*, 48, 240–267.
- Roher, B. (2012). *Enjeux et stratégies pour le futur de la filière abricot en Valais*.
- Roper, S., Vahter, P., & Love, J. H. (2013). Externalities of openness in innovation. *Research Policy*, 42(9), 1544–1554.
- Rossi, F., & Rosli, A. (2013). *Indicators of university-industry knowledge transfer performance and their implications for universities: Evidence from the UK's HE-BCI survey* (No. 13).
- Smith, W. R. (1995). Product Differentiation and Market Segmentation as Alternative Marketing Strategies. *Journal of Marketing*, 4(3), 63–65.
- Valais-Wallis.Promotion. (2015). Communiqué de presse. Saison des abricots 2015. *IFELV, Valais-Wallis Promotion*, 41(0), 17–19.



Interdisciplinary and transdisciplinary institutions: do they constitute peripheries among cultures?¹

Bianca Vienni* and Ulli Vilsmaier**

*Bianca.Vienni@leuphana.de

Center of Methods, Leuphana University of Lüneburg, Scharnhorststr. 1, C4.008, 21335 (Lüneburg, Germany)
Academic Department, Espacio Interdisciplinario, Universidad de la República (Uruguay)

** vilsmaier@leuphana.de

Center of Methods, Leuphana University of Lüneburg, Scharnhorststr. 1, C4.008, 21335 (Lüneburg, Germany)

ABSTRACT

The paper presents the progress done in the project entitled “Challenges in interdisciplinary and transdisciplinary knowledge production: institutions, cultures and communities”. This research investigates challenges of interdisciplinary (ID) and transdisciplinary (TD) knowledge production, focusing on processes of institutionalization, cultural transformations and the characteristics of communities.

The starting points for this research are two universities that have tackled the challenge of incorporating ID and TD in their institutional structure and study programs: the Center of Methods (Leuphana University Lüneburg, Germany) and the Espacio Interdisciplinario (Universidad de la República, Uruguay) (UdelaR).

The conclusions are not closed only to Germany and Uruguay – but compared to other examples to draw a general model to describe answer the question addressed in this paper as well as to assess ID and TD institutions and to systematize learning in terms of fostering and hindering factors for ID institutionalization. This seeks to epistemic cultures that allow a defined identity based on those features of ID and TD evaluated as general enough to be valid in the German and Uruguayan academic contexts and not only in peripheries. In this sense, we want to contribute to the methodological and theoretical construction of what it can be named “Studies on Interdisciplinarity” in Uruguay and Latin America.

INTRODUCTION

The paper presents the progress done in the project entitled “Challenges in interdisciplinary and transdisciplinary knowledge production: institutions, cultures and communities”. This research investigates challenges of interdisciplinary (ID) and transdisciplinary (TD) knowledge production, focusing on processes of institutionalization, cultural transformations and the characteristics of communities.

The specific objectives include: (i) historical analysis of the background of ID and TD knowledge production and the societal context of two universities, (ii) to analyse the current state of ID and TD knowledge production with the focus on epistemic living spaces (Felt 2009) and epistemic communities as framing concepts for the analysis and (iii) to put them

¹ This work is supported by Center of Methods (Leuphana Universität, Lüneburg) and Council of Scientific Research (Universidad de la República, Uruguay).

into value to strengthen future university policies in the context of these institutions and to develop transformative strategies that enable ID and TD development beyond structural barriers with a particular focus on key stakeholders and relevant actors.

The starting points for this research are two universities that have tackled the challenge of incorporating ID and TD in their institutional structure and study programs: the Center of Methods (Leuphana University Lüneburg, Germany) and the Espacio Interdisciplinario (Universidad de la República, Uruguay) (UdelaR).

This paper presents the partial results of objective (i): the historical analysis of the institutional background taking into account the concept of ‘epistemic cultures’ (Felt, 2009) and ‘peripheries’. The main questions addressed here are: (i) Did the centres that serve as case studies historically constitute peripheries among (epistemic) cultures and if so why? (ii) Is the concept of “periphery” a valid indicator to address these two ID and TD institutions? And (iii) Do those institutions represent peripheries and if so in which sense?

RATIONALE

We agree with the definition applied by Paasi (2013) to regions where institutionalization is a process in which the territorial (fixed or fuzzy borders, in this case disciplinary borders), symbolic (name and more material symbols) and institutional ‘shapes’ of an organizational form emerge, thereby forging the institution as an established, typically administrative unit (with defined frontiers) in the wider educational system and societal consciousness. The shapes are abstractions that help make sense of and elucidate the historically contingent, context-bound power relations embedded in complex institutional practices (governance, politics, economy, and media, for example) which ‘mediate’ ID and TD territories.

Secondly, culture refers to praxis carried out by certain communities and includes: norms, interests, conflicts, values, perceived by researchers, stakeholders and other actors involved in institutions (Strathern, 2004). They become ‘epistemic cultures’, defined as a specific group of social actors who share a belief in a common set of relationships and common values according (Felt, 2009).

In consequence, in ID or TD research knowledge production can be considered as an intercultural endeavour (Vilsmair et al, 2015). ID and TD research are particularly challenging as different epistemic communities (e.g., disciplines, knowledge fields) and fields of practice (e.g., civil society groups, administration, among others) are involved in the knowledge production.

BACKGROUND

Leuphana University of Lüneburg (Germany) is defined as a public university for the civil society of the 21st century. It seeks to establish a culture of learning through research within the structure of a modern democratic university (Spoun and Kölzer, 2014). The Center for Methods is an interfaculty institution within Leuphana University. The objective of the center is to provide support for students on all levels (Bachelor, Master, and PhD) with respect to education and training in methods and to offer different kinds of courses for ID, TD and un-disciplined methods. It also aims to support researchers across all disciplines in the application and development of scientific methods and contributes to create an ID and TD campus culture through joint research, events and team-teaching.

For its part, Universidad de la República (Uruguay) is the leading state institution for Higher Education and research in Uruguay. In collaboration with a wide range of institutional and social actors, it also conducts various activities aimed at developing socially valuable use of knowledge and the dissemination of culture (Arocena, 2008). Since 2008, the Espacio Interdisciplinario (EI) promotes collaboration among different disciplines to provide an integrated approach to problems which nature is not contained within the boundaries of individual disciplinary areas. It was founded as part of a University Reform undertaken by UdelaR in the period 2006 - 2014. The EI has to develop the three university objectives of teaching, research and community engagement.

The two cases of study were chosen as they present a set of characteristics which make them appropriate for this kind of analysis. At Leuphana University changes that indicate a move toward more sustained interdisciplinary structures can already be observed. This aim is currently being discussed by UdelaR as well.

METHODOLOGY

The approach guiding this study is in itself interdisciplinary and was generated from pre-existing findings in the literature mainly on the Science, Technology and Society Studies (STS) (Felt et al., 2015) and the so-called “Studies on Interdisciplinarity” (Darbellay, 2015).

In a previous study (Vienni, 2014) a methodological strategy was specifically designed in order to cope with a similar research problem. This methodological strategy, named “inter-approach”, evaluates the potential of an interdisciplinary approach on the basis that the combination, integration and convergence of different bodies of knowledge seek to collaborate in the construction of an enriching research process.

The comparative perspective this research implies, requires the integration of temporality concerning the socio – historical context when the processes of ID and TD knowledge production, the transformation of epistemic cultures and fields of practice occurred. We take into consideration the specific tasks of a university (research, teaching, transfer and outreach activities understood as a bidirectional process) as we consider these shape the process of knowledge production and ID and TD practices. Another relevant factor is that the two institutions taken as case studies are considered as “centres” in the sense of outstanding innovations in Higher Education. Nevertheless, ID and TD knowledge production still constitutes the “periphery” in the scientific rationale. In this sense, we may consider these institutions as constituents of “the centrality in the periphery”.

The historic analysis implies:

1. Construction of theoretical frame (concepts of periphery, centre, etc.).
2. Systematization of internal documents, reports and resolutions related to ID and TD centers at Leuphana University and UdelaR.
3. Integration of the dimensions of analysis through an inter – approach strategy (Vienni, 2014). The exposure of a system of relationships not only seeks to interpret reality in terms of an analytical model defined in advance, but uses a flexible methodological design that allows for changes in the course of the investigation.

RESULTS

The conclusions are not closed only to Germany and Uruguay – but compared to other examples to draw a general model to describe answer the question addressed in this paper as well as to assess ID and TD institutions and to systematize learning in terms of fostering and hindering factors for ID institutionalization. This seeks to epistemic cultures that allow a defined identity based on those features of ID and TD evaluated as general enough to be valid in the German and Uruguayan academic contexts and not only in peripheries. In this sense, we want to contribute to the methodological and theoretical construction of what it can be named “Studies on Interdisciplinarity” in Uruguay and Latin America.

REFERENCES AND CITATIONS

- Arocena, R. 2008. Presentación. In *Espacio Interdisciplinario, Jornada de Presentación*, Equipo de Rectorado de la Universidad de la República (ed.), Universidad de la República, Montevideo. Digital document in www.ei.ur.edu.uy.
- Darbellay, F. 2015. Rethinking inter- and transdisciplinarity: Undisciplined knowledge and the emergence of a new thought style. *Futures* 65 (2015), pp. 163 – 174.
- Felt, U. 2009. Knowing and living in academic research. In: Felt, U. (Ed.) *Knowing and living in academic research: convergence and heterogeneities in European research cultures*. Prague. Institute of Sociology of the Academy of Science of the Czech Republic.
- Felt, U., J. Igelsböck, A. Schikowitz and T. Völker 2015. *Transdisciplinary sustainability research in practice: between imaginaries of collective experimentation and entrenched academic value orders*, Preprint. Published by the Department of Science and Technology Studies, University of Vienna, May 2015, Available at sts.univie.ac.at
- Frodeman, R. 2014. “The end of disciplinarity”. In P. Weingart y B. Padberg (editors), *University experiments in Interdisciplinarity. Obstacles and Opportunities*, Bielefeld, Transcript, Science Studies, pp. 175 - 198.
- Paasi, Anssi 2013. Fennia: positioning a “peripheral” but international journal under conditions of academic capitalism. *Fennia* 191:1, pp. 1 – 13.
- Spoun, S. and C. Kölzer 2014. A place apart: Opportunities in developing Leuphana University of Lüneburg. In P. Weingart y B. Padberg (eds.) *University experiments in Interdisciplinarity. Obstacles and Opportunities*. Transcript, Science Studies, Bielefeld, pp.81-94.
- Strathern, M. 2004. *Common and bordelands. Working papers on interdisciplinarity, accountability and the flow of knowledge*, Wantage, Sean Kingston Publishing.
- Vienni, B. 2014. Interdisciplinary Socialization of Archaeological Heritage in Uruguay, *Journal of Cultural Heritage Management and Sustainable Development*, Vol. 4 No. 1, pp. 95-106.

Vilsmaier, U., M. Engbers, P. Luthardt, R. Maas – Deipenbrock, S. Wunderlich and R. Scholz 2015. Case-based Mutual Learning Sessions: knowledge integration and transfer in transdisciplinary processes. *Sustain Sci* (2015) 10:563–580.

Weingart, P. 2014. “Interdisciplinarity and the new governance of universities”. In P. Weingart y B. Padberg (eds.) *University experiments in Interdisciplinarity. Obstacles and Opportunities*. Transcript, Science Studies, Bielefeld, pp. 151 – 174.



"Putting in more than you take out" Towards evaluating research based on its public (not private) contributions¹

Paul Benneworth*, Julia Olmos-Peñuela** and Elena Castro-Martínez***

* p.benneworth@utwente.nl

Center for Higher Education Policy Studies (CHEPS), University of Twente, University of Twente, PO Box 217, 7500AE Enschede, The Netherland.

**julia.olmos@uv.es

Department of Management, Faculty of Economics, University of Valencia, Avenida Tarongers, s/n, 46022, Valencia, Spain.

***ecastrom@ingenio.upv.es

INGENIO (CSIC-UPV), Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia, Spain.

Abstract

There has been an increasing interest within research policy to measure the impact that research had in society, but also among academics to understand how research creates impacts beyond the traditional measures of patent licenses and spin-off creation. In these debates, there is often an emergent gap between research impact conceptualisations of ideas creating capacity and change in society, and the reality that it is very hard to measure the flow of ideas, but rather easy to measure transactions linked with those flows. In this paper we are concerned with the indicator periphery that emerges as a consequence of the impacts of some kinds of research better fitting to these underlying transactions than others. A range of lacunae emerge in indicators where the impacts of particular fields are badly captured by transactional measures – e.g. in social science and humanities (SSH) disciplines. This study addresses “how can we typologise the non-transactional ways in which publically funded research creates public benefits?” We start from a sociological approach to consider academic contribution to societal changes as the rate of inflow of usable knowledge into a reservoir (i.e. latent potential knowledge accumulated to later be exploited) dependent on the extent to which academic knowledge is cognate with potential social users. We draw on the openness framework to consider the five micro-practices by which scientists make research cognate with potential users, and we seek to categorise these micro-practices according to the different kinds of situated social learning practices they embody to create specific knowledge. To do so, we draw on twelve cases of study of Spanish SSH research groups (within the Siampi project) to develop a detailed typology of the kinds of micro-practices associated with openness as the basis for a new perspective for indicators of research impact that goes beyond the traditional transactional ones.

¹ This work was supported by the European Commission 7th Framework Programme (Social Impact Assessment Methods for research and funding instruments through the study of Productive Interactions between science and society - SIAMPI).

Introduction

There has in the last decade or so been an increasing driven within research policy to measure the impact that research had in society (Derrick & Samuel, 2016). We see this in a range of European countries, where research evaluation systems either account for societal impact in various different ways (UK, Netherlands), measure research impact through standing surveys (Spain, the UK) or have attempted to develop robust measures of research impact (Netherlands, Sweden). Although early attempts to measure research impact often followed direct impacts through translating transactions to national levels through equilibrium models, dissatisfaction with these more quantitative approaches has seen the rise of more qualitative approaches, typified by the UK's Research Excellence Framework, in which expert peer panels make qualitative judgements regarding the scope, depths and excellence of research impacts.

There is also increasing academic interest in understanding how research creates impacts beyond licensing deals, spin-off companies and patents, often with reference to these policy interventions. In these debates, there is often an emergent gap between research impact conceptualisations of ideas creating capacity and change in society, and the reality that it is very hard to measure the flow of ideas, but rather easy to measure transactions which may associate with those flows (Benneworth, Hazelkorn, & Gulbrandsen, 2016). In this paper, we are concerned with one aspect of this, the indicator periphery that emerges as a consequence of the impacts of some kinds of research better fitting to these underlying transactions than others. A range of lacunae emerge in indicators where the impacts of particular disciplines and fields are badly captured by transactional measures. Olmos-Peñuela, Benneworth, & Castro-Martínez (2014) for example demonstrate that in particular in the humanities and social sciences, there is a much greater propensity towards informal knowledge interactions than in many of the hard sciences. Our overarching research question in this research in progress is “how can we typologise the non-transactional ways in which publically funded research creates public benefits?”

Our diagnosis here is that this issue arises because of the implicit adoption of the transaction-equilibrium effect heuristic for research achieving impact. To provide an alternative perspective, we turn to sociological rather than economic perspective of science studies. We observe that in this tradition, useful knowledge represents a kind of reservoir of latent potential to later be exploited (in the tradition of the 1970 TRACES study). We therefore propose to consider academic contribution to societal capacity as the rate of inflow of usable knowledge into this reservoir (Hanney, Gonzalez-Block, Buxton, & Kogan, 2003), an additive knowledge production (Becher & Trowler, 2001; Latour & Woolgar, 1979) in turn dependent on the extent to which academic knowledge is cognate with potential social users. We in turn consider the micro-practices by which academics make their research cognate with potential users, drawing on the Openness Framework we have developed elsewhere (Olmos-Peñuela, Benneworth, & Castro-Martínez, 2015, 2016). Drawing on a detailed study of Spanish research groups in the social sciences and humanities within the Siampi project, we develop a detailed typology of the kinds of micro-practices associated with openness as the basis for a new perspective for indicators of research impact.

From an individual to collective perspective on research impact

Current research impact approaches demonstrate a mismatch – on the one hand there is a widespread appreciation of the different kinds of ways in which research creates value in society, but tempered on the other by a tendency in various ways to reduce this to things that

we all know are poor proxies for that impact. That is a severe public policy problem, because if we measure and stimulate the wrong things, we are steering the system towards working to produce the wrong outputs and ultimately creating the opportunities for public value failures. The great power of the economic model is that it is scalable – it takes a micro-level phenomenon and converts it into an aggregate effect that can be compared with other very dissimilar things (Benneworth, 2015; Benneworth, Hazelkorn, & Gulbrandsen, 2016), something which policy-makers find very useful (Molas-Gallart, 2015).

We take a slightly different perspective here, in that we note that in the science studies literature there is a much less instrumentalist stream that questions the extent to which that knowledge exchange comes through immediate researcher-user interactions (Sarewitz & Pielke Jr., 2007). Perhaps fitting with Louis Pasteur's idea of the 'prepared mind' (Garud, Kumaraswamy, & Karnøe, 2010), useable knowledge can be considered as constituting a reservoir which can then later be exploited (Sarewitz & Pielke Jr., 2007). Those interested in *innovation* policy should seek to maximise the outflow of the knowledge from that pool, at which point it becomes transformed and embedded in particular artefacts to which property rights apply. But if you are interested in *science* policy, and in particular publicly funded science, then what is critical here is the rate of **inflow** of knowledge into the pool, from the base of publicly funded science.

Our contribution to this debate about usefulness in science policy is our relation of the rate of inflow to the idea of **user cognateness**; the reservoir metaphor captures all knowledge that might later be taken up by a user. We therefore argue that a characteristic of such knowledge that may later flow out of the 'reservoir' is that it has a cognateness with users, because that cognateness is the basis for any kind of knowledge exchange (Boschma, 2005; Fromhold-Eisebith, Werker, & Vojnic, 2014). We further argue that newly created knowledge is cognate with the knowledge upon which it is build. From this, we deduce that knowledge that is created using 'user knowledge' in some way will have a cognateness later allowing users to exploit it. We therefore argue that (in the frame of our metaphor) the rate of flow into the knowledge reservoir is associated with the extent to which those creating that knowledge incorporate user knowledge into their research micro-practice (our 'openness' variable).

To date we have argued that there are five overarching kinds of research micro-practice where knowledge is materially combined to create new knowledge, and thus where user knowledge may be involved (Olmos-Peñuela et al., 2015). But this approach remains very broad in its perspective, grouping activities that whilst conceptually very similar (research question design, for example), do vary considerably in practices between different fields of study. We are concerned that behind these micro-practices are social learning behaviours in which researchers co-create knowledge in various ways involving users and therefore making it cognate with a wider set of users. Therefore, in this paper we focus on the detail of individual social learning behaviours by which researchers incorporate user knowledge.

For our conceptual model of social learning behaviours, we here use the Amin & Roberts (2008) & Roberts (2014) typologies of different kinds of situated learning practices. Scientific knowledge may be unsituated and generic/ universal, but if there are situated learning processes, then it is the kind of research that can later be fixed and transformed into useful knowledge (Vohora, Wright, & Lockett, 2004). So in this paper we categorise the different research micro-practices according to the different kinds of situated social learning practices they embody to create specific knowledge.

Research progress to date

To address this we will draw on a set of interviews undertaken within the Siampi project (addressing the Social Impact Assessment Methods for research and funding instruments through the study of Productive Interactions)² to explore the social learning behaviours associated with the different research micro-practices, to develop a more thorough understanding of how researcher groups build up the cognateness and thus usability of their research. The Siampi dataset was generated under the auspices of a larger European research project involving case studies from four countries in the period 2009-2011, seeking to provide a deeper qualitative understanding of scientific and societal interactions within science. Although we are repurposing the Siampi data for our own needs, the focus of the interviews, there is sufficient correspondence between the questionnaire structure and the openness model to make the dataset fit for purpose, and as an exploratory piece of work we further argue it saves the need for specific activity.

In the Spanish dataset, there are twelve interviews with Social Sciences and Humanities (SSH) research groups where interviews were undertaken with these research groups, as well as with users of the knowledge created by that group. We will seek firstly to identify whether there was evidence in each of the research micro-practices of shared social learning with those partners, and in the case that there was, to classify it according to the Roberts' framework. On that basis we will be able to identify the relative diversity of social learning practices within those research micro-practices by which openness is created as the basis for identifying alternative impact indicators.

The expected output of this research will be materialised in a typology of the kinds of micro-practices associated with openness that may set the basis for a new approach for research impact indicators that goes beyond the traditional transactional ones.

References

- Amin, A., & Roberts, J. (2008). Knowing in action: Beyond communities of practice. *Research Policy*, 37, 353-369.
- Becher, T., & Trowler, P. R. (2001). *Academic tribes and territories: Intellectual enquiry and the culture of disciplines*. Buckingham: The Society for Research into Higher Education & Open University Press.
- Benneworth, P. (2015). Tracing how arts and humanities research translates, circulates and consolidates in society. How have scholars been reacting to diverse impact and public value agendas? *Arts and Humanities in Higher Education*, 14, 45-60.
- Benneworth, P., Hazelkorn, E., & Gulbrandsen, M. (2016). *The impacts and future of arts and humanities research*. London: Palgrave (in press).
- Boschma, R. (2005). Proximity and innovation: a critical assessment. *Regional Studies*, 39, 61-74.
- Derrick, G. E., & Samuel, G. N. (2016). The Evaluation Scale: Exploring Decisions About Societal Impact in Peer Review Panels. *Minerva*, 54, 75-97.

² <http://www.siampi.eu/>

- Fromhold-Eisebith, M., Werker, C., & Vojnic, M. (2014). Tracing the social dimension in innovation networks. In R. Rutten, P. Benneworth, F. Boekema & D. Irawati (Eds.), *The Social Dynamics of Innovation Networks* (pp. 221-239). London: Routledge.
- Garud, R., Kumaraswamy, A., & Karnøe, P. (2010). Path dependence or path creation? *Journal of Management Studies*, 47, 760-774.
- Hanney, S. R., Gonzalez-Block, M. A., Buxton, M. J., & Kogan, M. (2003). The utilisation of health research in policy-making: concepts, examples and methods of assessment. *Health Research Policy and Systems*, 1: 2, <<http://health-policy-systems.biomedcentral.com/articles/10.1186/1478-4505-1-2>> accessed 11 February 2016.
- Latour, B., & Woolgar, S. (1979). *Laboratory life: The social construction of scientific facts*. London: Sage.
- Molas-Gallart, J. (2015). Research evaluation and the assessment of public value. *Arts and Humanities in Higher Education*, 14, 111-126.
- Olmos-Peñuela, J., Benneworth, P., & Castro-Martínez, E. (2014). Are ‘STEM from Mars and SSH from Venus’?: Challenging disciplinary stereotypes of research’s social value. *Science and Public Policy*, 41, 384-400.
- Olmos-Peñuela, J., Benneworth, P., & Castro-Martínez, E. (2015). What stimulates researchers to make their research usable? Towards an ‘openness. *Minerva*, 53, 381-410.
- Olmos-Peñuela, J., Benneworth, P., & Castro-Martínez, E. (2016). Does it take two to tango? Factors related to the ease of societal uptake of scientific knowledge. *Science and Public Policy*, (forthcoming).
- Roberts, J. (2014). Community and the Dynamics of Spatially Distributed Knowledge Production: The Case of Wikipedia. In R. Rutten, P. Benneworth, D. Irawati & F. Boekema (Eds.), *The social dynamics of innovation networks*. London: Routledge.
- Sarewitz, D., & Pielke Jr., R. A. (2007). The neglected heart of science policy: reconciling supply of and demand for science. *Environmental Science & Policy*, 10, 5-16.
- Vohora, A., Wright, M., & Lockett, A. (2004). Critical junctures in the development of university high-tech spinout companies. *Research Policy*, 33, 147-175.



On multiplying methods in the field of research evaluation

Gemma Derrick (Lancaster University, g.derrick@lancaster.ac.uk)

Jordi Molas-Gallart (INGENIO & SPRU, jormoga@ingenio.upv.es)

Sarah de Rijcke (CWTS, Leiden University, s.de.rijcke@cwts.leidenuniv.nl)

Ingeborg Meijer (CWTS, Leiden University, i.meijer@cwts.leidenuniv.nl)

Inge van der Weijden (CWTS, Leiden University, i.c.m.van.der.weijden@cwts.leidenuniv.nl)

Paul Wouters (CWTS, Leiden University, p.f.wouters@cwts.leidenuniv.nl)

ABSTRACT

This special session forms part of a larger program aimed at the multiplication and integration of methodological approaches in the research evaluation and innovation policy field. The session builds on previous initiatives by Gemma Derrick and colleagues at CWTS, INGENIO, the Rathenau Instituut and SPRU, exploring the advantages of qualitative methodological tools at the STI/ENID conference in Lugano, and an international workshop in London in October 2015. The program is highly topical: the research evaluation field is currently reconsidering its methodological foundations in light of new research questions arising from policy initiatives regarding a) the move toward open science; b) a reconceptualization of research excellence to include societal relevance; c) diversification of academic careers, and d) the search for indicators showcasing responsible research behavior and innovation. This new special session at STI2016 will advance and broaden the scope of previous initiatives by *building bridges* between cutting edge research involving quantitative, qualitative, and mixed methodological research designs. Bringing together leading experts and promising researchers with distinctive methodological skill-sets, the session will demonstrate the advantages of cross-fertilization between ‘core’ and ‘peripheral’ methodological approaches for the research evaluation and science indicators field.

PROGRAM FOR THE SESSION

1. Introduction (10 min):

The session starts with a brief introduction by Gemma Derrick (Lancaster University) to recap the major highlights of the special session on Qualitative and Mixed Methods (QMM) at the STI/ENID conference in Lugano (2015) and the QMM workshop at Brunel, funded by the British Academy Rising Star Engagement Award program. The 2015 session and workshop centred on discussing the value of qualitative research methods, and on showcasing current research involving qualitative or mixed methodological research designs in the field of research evaluation.

2. Provocation (10 min):

In this second part of the special session, Paul Wouters (CWTS, Leiden University) will discuss the rise of mixed methods research, and ask what it actually means to mix methods in measuring research quality. His contribution ends with a provocation to think ‘Beyond mixed method’.

3. Presentations/‘Demonstrations’ (60 min; 20 min for each pair):

The third part of the session is organised around ‘think pairs’, linking three presenters to three respondents. The presenters will be asked to prepare a contribution that showcases their methodological approach and base their presentation on on-going or recently completed work. There will be one presentation centred on a mainly quantitative approach, one presentation based on mixed-methods approaches, and one presentation that primarily utilises qualitative methods. Each presentation will be juxtaposed by a dedicated response from scholars with complementary methodological expertise. Respondents are asked to ‘think with’ the presenter about the role of the methods in the projects, in relation to the research questions, empirical material/data, and results. The composition of the think pairs will be decided closer to the date of the conference.

The primary purpose of the session is to encourage methodological cross-fertilization and develop ideas for future research. A further aim is that, through presentations and related discussion, participants will work together to experience the mutual benefits of mixed methodological studies in the field.

4. Next steps (10 min)

The session will end with a brief update on the institutional embedding and future activities for the international QMM network in STI by Gemma Derrick, Jordi Molas-Gallart (Ingenio) and Sarah de Rijcke (CWTS). The main objectives of the QMM network are to promote the multiplication of methodological approaches in the research evaluation field; encourage standards of excellence and good practice for mixed methods for evaluation research; engage a dialogue that contributes to a method-led reflection of the evaluation field; provide opportunities for the open discussion and development of the above goals through researcher and practitioner interaction. The network currently receives support from the British Academy. We are currently putting together a special issue for Research Evaluation on the topic, and the session will be used as a platform to encourage submissions to further the network and the broader initiative.

The session will be chaired by Inge van der Weijden (CWTS).

Intended audience

Researchers within the field of research evaluation who have mainly used either quantitative or qualitative approaches in their current research, and those who are interested in adopting new approaches in their future research; current and future members of the QMM network.

Length

90 minutes

Preferred number of participants

30-40

Special requests/equipment needs

Beamer/laptop

Flip-over



Why DORA does not stand a chance in the biosciences¹

Jochen Gläser*

*Jochen.Glaser@ztg.tu-berlin.de

Center for Technology and Society, TU Berlin, Hardenbergstr. 16-18, 10623 Berlin (Germany)

ABSTRACT

The San Francisco Declaration on Research Assessment (DORA) has become emblematic for the international movement against the use of journal impact factors (JIFs) in research assessment. This movement gained an institutional stronghold in Australia when in 2010, i.e. even before DORA, the Australian National Health and Medical research Council decided to remove JIFs from their assessment of grant proposals. Given the strong dependency of Australian researchers on funding from the research councils, one could expect the decisions by the funding councils to have an impact on the use of JIFs in individual research assessment. The aim of this paper is to demonstrate that there was no impact beyond a change of language in assessments and justifications of decisions, and to explain why this is the case. Interviews with officers of the two funding councils, university managers and researchers in cell biology and biostatistics were conducted in Australia in March 2016. The interviews explored the use of impact factors in seven roles of a biomedical researcher's role-set, namely reader, author, group leader, grant applicant, reviewer of grant applications, university employee and university manager. The JIF solved three major problems for incumbents of the various roles: it reduced complexity, increased efficiency, and compensated for unfamiliarity in cases in which assessments were required but could not be conducted by judging content. Its use has not been changed by the research council's intervention. JIFs are unlikely to be overcome as tools for research assessment in the life sciences unless another tool that solves the same problems is offered.

INTRODUCTION

The San Francisco Declaration on Research Assessment (DORA) has become emblematic for the international movement against the use of journal impact factors (JIFs) in research assessment. Its general recommendation reads as follows:

“Do not use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist's contributions, or in hiring, promotion, or funding decisions.” (DORA 2013: 2)

This movement gained an institutional stronghold in Australia when in 2010, i.e. even before DORA, the Australian National Health and Medical research Council (NHMRC) decided to remove journal impact factors from their assessment of grant proposals (NHMRC 2010). The Australian Research Council soon followed suit. Applicants could no longer submit JIFs on their lists of publications, and specialist reviewers of proposals or panel members were not allowed to use JIFs in reviews or decisions on the funding of a proposal. Given the strong dependency of Australian researchers on funding from the research councils, one could expect

¹ This work was supported by the German Federal Ministry of Education and Research (Grant 01PY13015).

the decisions by the funding councils to have an impact on the use of JIFs in individual research assessment.

The aim of this paper is to demonstrate that there was no impact beyond a change of language in assessments and justifications of decisions, and to explain why this is the case. Previous self-observations of biomedical scientists and ethnographic research have highlighted the use of JIFs by researchers themselves in their decisions on publication channels (Rushforth & de Rijcke 2015). Rushforth and de Rijcke also mention the importance of JIFs for getting grants and prevailing in intra-organisational evaluations at universities (ibid.). However, these accounts are largely descriptive, and are unable to explain why JIFs are such an incredibly stable aspect of research practices in the biomedical sciences.

APPROACH

Explaining why researchers continue certain practices despite an intervention by influential actors requires identifying the mechanisms that stabilize the use of JIFs and the conditions under which they operate. This can be achieved by systematically exploring the use of impact factors in the various roles of a biomedical researcher's role-set (Merton 1957 on this concept). If we apply this idea to the position of a biomedical researcher at a university, seven roles in which the JIF potentially plays a role can be analytically distinguished (Table 1). In each of these roles, researchers may use the JIF, and will experience feedback on their use of JIFs based on the success of their actions. These feedbacks will contribute to the development of researchers' subjective theories. Through being informed by and contributing to the frames, the uses of JIFs in each role stabilize each other.

Interviews with officers of the two funding councils (7), university managers (3) and researchers in cell biology (4) and biostatistics (2) were conducted in Australia in March 2016. The number of interviewees with whom each role was explored (Table 1) exceeds the number of interviewed academics because some managers could also be interviewed about recent or current researcher roles. Since responses were consistent across cases, 'saturation' (Glaser & Strauss 1967: 61-62) could be achieved with a relatively small number of interviews.

Table 1. Number of instances that were explored for each role in an academic's role set.

Role	Practices in which JIFs are used	Number of interviewees with whom role was explored
Reader	searching for literature, selecting literature for reading, reading	8
Author	selecting a journal for publication and publishing an article	8
Research group leader	recruiting researchers, planning experiments	5
Grant applicant	writing, submitting and negotiating grant applications	7
Assessor of grant proposals	assessing proposals and applicants, making decisions on which grant applications to fund	8
University employee	complying with the university's	5

	expectations, applying for promotion	
University manager	formulating and communicating performance expectations, assessing performance and responding to assessments	4

Interviews were preliminarily analysed according to the functions of JIFs for each of the roles, feedback researchers and managers receive for their use of JIFs, and resulting contributions to the (in)stability of JIF use. The following results are based on this preliminary analysis.

PRELIMINARY RESULTS

Table 2 lists the functions the JIF has in each of the roles, feedbacks experienced by researchers, and resulting contributions to researchers' subjective theories about the JIF. The JIF solved three major problems for incumbents of the various roles: it reduced complexity, increased efficiency, and compensated for unfamiliarity in cases in which assessments were required but could not be conducted by judging content.

Table 2. Functions of JIFs, positive feedback received from their use and contributions to subjective theories about JIFs

Role	Function of JIF	Positive feedback	Contribution to subjective theory
Reader	Reducing complexity of article selection	Collection of relevant information	Important work (explanatory accounts backed by sufficient empirical evidence) is published in journals with high impact factors, imperfect correlation
Author	Reducing complexity of journal selection, accelerating publication, reaching target audience, maximising reputation	Increasing reputation	Important work (explanatory accounts backed by sufficient empirical evidence) is published in journals with high impact factors, imperfect correlation
Research group leader	Informing decisions on experiments	Efficient research	Journals with high JIFs demand more data and mechanisms that explain them
Grant applicant	Demonstrating reputation	Successful application	JIF is used as indicator of research quality
Assessor of grant proposals	Increasing efficiency, substituting for knowledge about content	Efficient assessment, consensus in decision-making	JIF can be used as indicator of quality and as substitute for content, is field-specific, should not be used in comparisons

University employee	Demonstrating reputation	Satisfaction of university management, promotion	JIF is used as indicator of research quality
University manager	Reducing complexity, substituting for knowledge about content	Efficient assessment, consensus in decision-making	JIF can be used as indicator of quality and as substitute for content, is field-specific, should not be used in comparisons

The resulting subjective theories consider the JIF as sufficiently reliable guide to the quality of publications in one field (comparisons across fields were considered impossible by all interviewees) regardless of the (universally acknowledged) cases of mismatch. In these theories, the actual numbers play a minor role. Instead, the JIF is used as a symbol of journal rank, and used as an indicator for ‘better’ and ‘worse’ journals. The most important stabilizing mechanism is ‘role-taking’ (Biddle 1986: 84-85). Researchers form assumptions about other researchers’ behavior by extrapolating from their own behavior and from observations. They base their reader behavior on assumptions about the behavior of authors and vice versa, their writing of grant applicants on assumptions about reviewer behavior. Employees form assumptions about managers’ behavior, and managers have been employees under evaluation themselves. Through role-taking, the use of JIFs in one role informs the uses of JIFs in the other roles, thereby ensuring consistency of JIF use and JIF interpretation in researchers’ subjective theories.

Interviewees from research councils confirmed that with the removal of the impact factor (and later the H-index), applicants asked what they should submit instead. The research councils tried to develop alternatives based on citations but this turned out not to be feasible. Researchers and research managers pointed out that citation counts have the disadvantage of not being able to provide current information on quality but can be applied only to work that has been published several years ago.

PRELIMINARY CONCLUSIONS

Three preliminary conclusions can be drawn from this study. First, JIFs are unlikely to be overcome as tools for research assessment in the life sciences unless another tool that solves the same problems is offered. The Australian research councils’ initiative remains ineffective because they failed to introduce an alternative to JIFs. Second, in the analysis of indicator use it is important to distinguish between endogenous indicators (those that are used in everyday research practices to solve recurrent problems) and exogenous indicators (those that solve problems only for management and science policy). The latter appear to be much more malleable but also have more potential to distort research practices. Third, more attention needs to be paid to the ‘loss of peeriness’, i.e. the rise of evaluation situations in which peer review is not peer review anymore due to increasing specialization in research.

REFERENCES

- Biddle, B. J. (1986). Recent Development in Role Theory. *Annual Review of Sociology*, 12, 67-92.
- DORA (2013). San Francisco Declaration on Research Assessment. Retrieved 29 March from: <http://www.ascb.org/dora/>
- Glaser, B. G. & Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine.
- Merton, R. K. (1957). The Role Set: Problems in Sociological Theory. *British Journal of Sociology*, 18, 106-120.
- National Health and Medical Research Council (NHMRC) (2010). NHMRC removes Journal Impact Factors from Peer Review of individual research grant and fellowship applications. Retrieved 29 March from: http://www.nhmrc.gov.au/_files_nhmrc/file/grants/peer/impact%20factors%20in%20peer%20review.pdf.
- Rushforth, A. & de Rijcke, S. (2015). Accounting for Impact? The Journal Impact Factor and the Making of Biomedical Research in the Netherlands. *Minerva*, 53, 117-139.



Are institutional missions aligned with journal-based or document-based disciplinary structures?

Richard Klavans* and Kevin W. Boyack**

* rklavans@mapofscience.com
SciTech Strategies, Inc., Wayne, PA 19087 (USA)

** kboyack@mapofscience.com
SciTech Strategies, Inc., Albuquerque, NM 87122 (USA)

ABSTRACT

Missions represent the underlying purpose of an institution. These missions can be focused (finding a cure for cancer) or diverse (providing all health services to a local population). They might be aimed at basic research (finding new sub-atomic particles) or very applied (forecasting tomorrow's weather). Missions can be extremely practical (building i-phones) or abstract (creating maps of scientific inquiry).

Our primary focus is on those institutions that are also contributing to society's knowledge about scientific and technical phenomena. The publications of these institutions are, to some degree, an implicit statement of their mission. Institutions focusing on a cure for cancer will publish articles associated with cancer, while hospitals will publish in a diverse set of medical specialties. Institutions focused on subatomic particles publish in specialized physics journals. While the publication profile of an institution is obviously not the same as an institution's mission, it is typically consistent with its mission.

In this study we analyze the publication profiles of over 4400 institutions using Scopus data to determine if their institutional missions are best explained using a journal-based classification system or a document-based classification system. The structure of this article is as follows. The background section places this work in the context of two streams of research – the accuracies of different document classification systems, and the effect of different national contexts (specifically wealth, health and democracy) on science systems and their impact. We then describe our data and methods before addressing two questions: Do the missions of certain types of institutions align with journal-based or article-based disciplines, and does this vary with national context (wealth, health and democracy). We conclude with a discussion of limitations and possible areas for further investigation.

BACKGROUND

Not all institutions have published mission statements. This is particularly true for very diverse institutions. For example, the largest institution in terms of publication activity, Harvard University, does not have a mission statement that covers its undergraduate college, graduate (including medical) schools, and research centers.¹ As one goes to more granular levels, however, missions become apparent. For example, the Harvard Medical School has 24 research centers focusing on such issues as AIDS, Cancer, and Translational Science.² The de facto missions of these research centers are to conduct leading research in their topic spaces.

¹ <http://www.harvard.edu/about-harvard/harvard-glance>

² <http://hms.harvard.edu/research/initiatives-programs-centers-and-institutes>

Institutions with a single focus may or may not have mission statements. For those that do, the mission statement is mirrored by their publication record. For example, the Japanese National Institute for Materials Science (NIMS) is a basic research laboratory with a focused mission. Its stated mission is to “achieve the policies of the Japanese government in the Science and Technology Basic Plan and its own 5-year Mid-term Program based thereon, and to contribute to materials research and the enhancement of science and technology”.³ NIMS is clearly focused on materials science, and its publications are predominantly in materials science journals. Overall, we suggest that *publication leadership is a reasonable statement of the mission of the institutions who are publishing in the scientific and technical literature.*

Why are the missions of institutions important? Why should we care? A great deal of effort goes into understanding and measuring science at the national level. Nations, however, are not the level at which science is performed. Rather, it is the institutions in each nation that ultimately bring national science strategies to fruition. The missions of the institutions combine to reflect the missions and strategies of nations. National strategies, missions, and publication profiles, and their correlation with wealth, health, and democracy have been studied (King, 2004; May, 1997; Moya-Anegón & Herrero-Solana, 2013). Institutional missions, however, have not.

The most well-known work on science and the wealth of nations are the landmark articles by May (1997) and King (2004). May’s (1997) article in *Science* suggested that publication profiles are influenced by national wealth. Using data from ISI, he focused mostly on the publication profiles (papers and citations) of the top 15 nations in 20 fields. National differences were viewed as areas of national strengths. For example, the UK was found to have greater strengths (higher than expected levels of publication and impact) in pharmacology, clinical medicine, plant and animal science, and neuroscience. May’s results were consistent with a contemporary study of the UK science base. He concluded with an interesting hypothesis that the differences in these national profiles might be reflected in the choice of institutional mechanisms in different nations.

King’s (2004) paper in *Nature* followed up on some of the same themes and extended the analysis to include the evaluation of a larger set (31) of nations. The database was the same (the ISI database) but the time period was more recent (1993-2001) and journal coverage was much higher (8000 journals vs. 4000 journals in May’s study). King points out that this sample covered 98% of all highly cited papers, while activities in the remaining 162 nations only accounted for 2%. He noted the relationship between overall citation impact and GDP per person (as his indicator of national wealth). He was able to replicate May’s observation that the UK has strengths (relatively more publication emphasis) in medicine and biology. King’s explanation about the different national profiles, however, was different. He focused more on stages of economic development (developed vs. emergent nations). He emphasized the importance of innovation (the specific indicator he mentions was R&D/GDP) but did not mirror May’s hypothesis that institutional structures (universities vs. research labs) played a role in the differences in national publication profiles.

There are a variety of less influential studies that have also focused on this issue. As an example, Doré et al. (1996) applied factor analysis to his dataset (48 nations, 17 fields) in

³ <http://www.nims.go.jp/eng/research/>

order to create a map that reflected different national strengths. The interpretations of their findings are as follows. Factor 1 made the distinction between disciplines associated with the industrial revolution (physics & chemistry) and the “more modern” disciplines associated with the life sciences (medicine, environment and computer science). The second factor dealt with nations that focused on agriculture vs. geoscience (the sea, earth’s mantle and mining). The third and fourth factors (which explain far less of the variance in the data) focused on smaller groups of disciplines.

Moya-Anegón & Herrero-Solana (2013) followed in the tradition of using factor analysis to detect the dimensionality of national strengths. The authors expanded the coverage of nations (94 countries) and fields (27 subject categories), finding only three factors. The interpretations of these factors were somewhat different than Doré. Factor 1 was more aligned with ‘health and democracy’ (democratic regimes invest more in medicine and correspondingly have longer life expectancies). Factor 2 is more aligned with regional issues (eastern Europe and Russia had a greater emphasis on physics, engineering and chemistry). Factor 3 was associated with third world nations.

This study follows in the same tradition but with the following differences. First, we argue that the unit of analysis should be the institutions in a nation rather than the nation. While the nation might have a unique set of strengths that are attributed to wealth, health and democracy, it is the institutions that, in pursuing their individual missions, produce national strengths. Aggregating all of the institutions into a national profile is, in essence, destroying all of the detailed information about the actors that might be pursuing different missions. We point out that there are no large scale empirical studies of the publication profiles of institutions in different fields and in different nations. To the best of our knowledge, this is the first study of this kind.

Second, previous work in this field has been exploratory, not confirmatory. Each of the studies mentioned above nominated different explanations, but because of the limitations in their sample size, could not test any of their claims. We therefore don’t know whether missions are associated with sectors (e.g., academia, industry, government), national wealth, R&D intensity or human development (an index that combines life expectancy, education and income per capita). We don’t know if these differences correlate with level of democracy or agricultural capabilities. By focusing on institutions (which increases the sample size to over 4400 actors), it is now possible to test many of the hypotheses raised in the literature.

The third major difference, which will be discussed more thoroughly in the methodology section, is the way in which fields (and thus missions) are defined. Fields are often defined in terms of journal-based disciplines, consistent with the way academic departments make promotion and tenure decisions. More recently, fields have been defined based on topics – groups of papers rather than journals – consistent with the way researchers self-organize around a problem with little regard for traditional disciplinary boundaries. We will test whether institutions and their missions conform more to journal-based or article-based structures. We will also test whether national context (e.g., wealth, health and democracy) has an effect on these results.

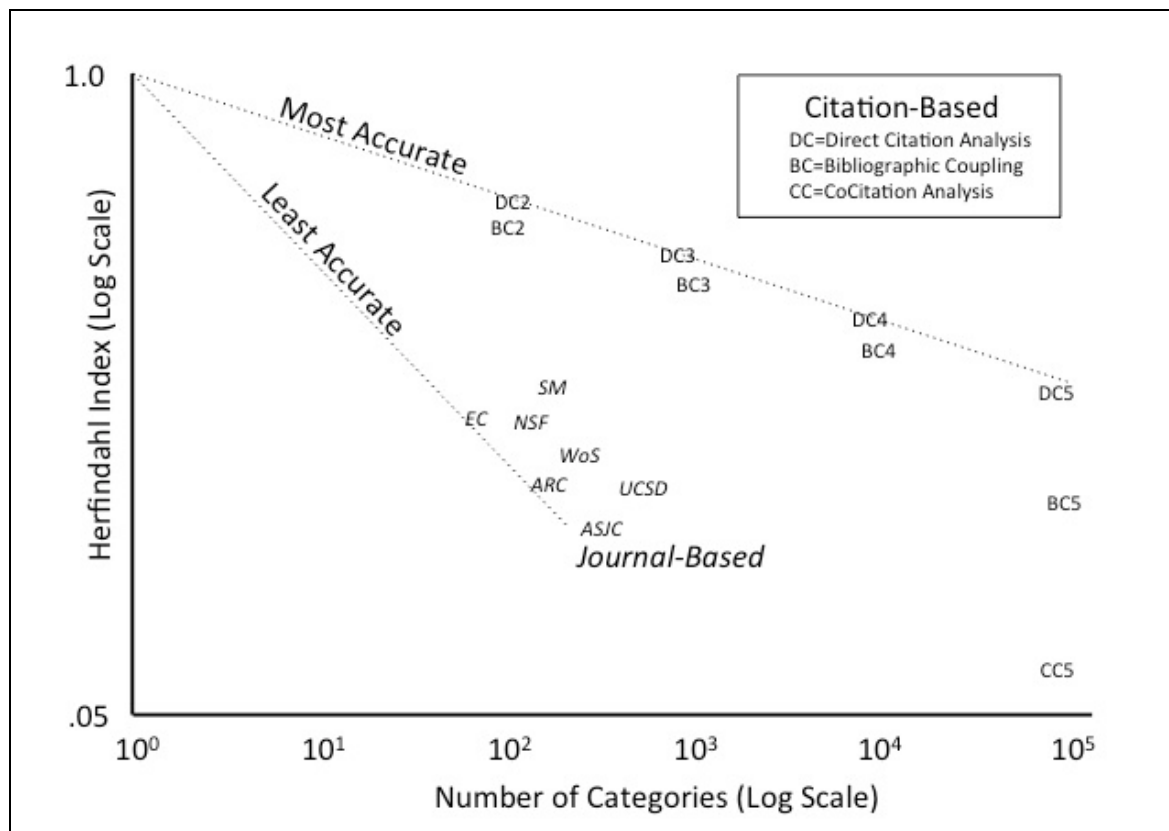
DATA AND METHODS

The publications of 4482 institutions were identified for the 5-year time period 2010-2014 using Scopus data. These institutions were identified manually several years ago by grouping

Scopus affiliation profiles into institutions for those institutions publishing at least 50 papers per year. Many institutions have only a single affiliation profile in Scopus. However, most of the larger institutions have medical (or other) schools or research institutes that, while they are clearly part of the parent institution, have separate affiliation profiles in Scopus. For example, our institutional profile for the University of California, San Diego (UCSD) contains six affiliation profiles including those for the Scripps Institution of Oceanography, UCSD School of Medicine, and San Diego Supercomputer Center. Each institution was also assigned to a sector (e.g., academia, industry, government), which was, in most cases, an obvious choice.

To answer the question as to whether institutional missions are more aligned with journal-based categories or article-based categories, we chose to compare three different classification systems, two of which are based on clustering of journals, and the other based on clustering of papers. These were chosen based on the results of our recent work that compared the accuracies of 16 different taxonomies (or classification systems) of science that were created using a variety of methods (Klavans & Boyack, 2016). That study included nine citation-based document classification systems (including direct citation, bibliographic coupling, and co-citation) and seven journal-based document classification systems. It analyzed the reference profiles in 37,000 gold standards (articles published in 2010 with 100 or more references), and calculated the Herfindahl index (the sum of the squared percentages by category) for each paper using all 16 classification systems. Figure 1 shows the results of that study, plotting the average Herfindahl value for each classification system as a function of the number of categories. The best possible value goes through the 1,1 point (1 category, $H=1.0$). Direct citation was shown to provide the most accurate taxonomy for all levels of granularity, from roughly 100 categories to 100,000 categories. All of the journal based classification systems (in italics) are far less accurate. The best journal-based approach is labeled SM (Science-Metrix). The worst citation-based approach is co-citation analysis. The worst journal-based approach is the ASJC classification system associated with the Scopus database (the classification system used by Moya-Anegón (2013)).

Our previous study clearly showed that paper-based classifications are far more consistent with the ways that authors cite papers than are journal-based classifications. Nevertheless, it is very possible that journal-based classifications, since they represent traditional disciplinary structures, might actually represent the publication profiles of institutions (and thus their missions) better than paper-based classifications. Given the results from Figure 1, and the possibility that journal-based methods are detecting a different phenomenon than paper-based methods, we decided to use the best document method (DC2) and best journal method (SM) where the number of categories are roughly the same (114 and 176, respectively) in this study. As a benchmark, we also use the NSF journal classification system (138 categories) to test whether it continues to generate less accurate results.

Figure 1: Relative accuracies of 16 clustering solutions (Klavans & Boyack, 2016).

We used the same method employed in our previous study to determine if the missions of institutions are more aligned with journal-based clusters or document-based clusters. Using the publication profiles of the 4482 institutions, Herfindahl indexes were calculated for each institution using the DC2, SM and NSF classification systems. We used the slope ($\log(\text{Herf})/\log(\#\text{categories})$) as our figure of merit. The slope is calculated by drawing a line from the approach (such as SM) to the 1,1 point (the hypothetical situation where there is one category and a Herfindahl of 1.0). All of the slopes will be negative (as shown in Figure 1) and the least negative curve represents the most accurate. The hypothetical line from the 1,1 point to SM shows that this approach is more accurate than any other journal-based approach, but less accurate than any of the document based approaches.

This procedure allowed us to determine if the mission of an institution was more aligned with journal-based clusters (the figure of merit would be higher for the NSF or SM classification system) or document-based clusters (the figure of merit would be higher for the DC2 classification system).

RESULTS

Missions: Journal-based vs. Document-based

Table 1 shows that, overall, almost two-thirds of the institutions in this sample are aligned with a document-based mission. The sector-based results are somewhat dramatic and unanticipated. Hospitals have journal-based missions, and seem to be much more ‘ivory towers’ than are universities. Universities are the most likely to follow document-based

missions, which are often interdisciplinary in nature. Industry is nearly as document-based as academia.

Table 1. Document-based missions by sector.

Sector	# Institutions	% Document-based missions
Hospital	559	14.3%
Academy of Science	120	60.0%
Government	724	62.6%
Non-profit	143	67.1%
Industry	297	71.7%
Academia	2629	76.1%
Not classified	10	90.0%
Total	4482	65.2%

National context

In order to evaluate the effect of national context, we collected nation-specific data on the following:

- wealth (2014 GDP per person / International Monetary Fund)⁴
- health (2015 Human Development Indicator / United Nations)⁵
- democracy (2015 Democracy index / Economic Intelligence Unit)⁶
- innovation (2010 R&D/GDP statistics / OECD)⁷
- agriculture (2011 hectare of arable land per capita / FAO)⁸

Correlations were calculated between these data using values for 83 nations. Table 2 illustrates that democracy, innovation and health are strongly inter-correlated. Innovation is linked to wealth. The indicator of agricultural focus of a nation is the most independent of the five contextual variables.

Table 2. Correlations between contextual variables (N=83 nations).

	Democracy	Innovation	Health	Wealth
Innovation	0.5650			
Health	0.5264	0.5917		
Wealth	0.0661	0.3635	0.1584	
Agriculture	0.0925	0.0429	0.1234	0.0551

We then conducted an ANOVA test using all 4482 institutions to determine which of these indicators is most likely to predict the sector associated with an institution. By far, the index of democracy had the highest predictive ability. Table 3 summarizes the effect on sectors.

⁴ [https://en.wikipedia.org/wiki/List_of_countries_by_GDP_\(nominal\)](https://en.wikipedia.org/wiki/List_of_countries_by_GDP_(nominal))

⁵ https://en.wikipedia.org/wiki/List_of_countries_by_Human_Development_Index

⁶ https://en.wikipedia.org/wiki/Democracy_Index

⁷ https://en.wikipedia.org/wiki/List_of_countries_by_research_and_development_spending

⁸ <http://www.nationmaster.com/country-info/stats/Agriculture/Arable-land/Hectares-per-capita#2011>

The data suggest that hospitals, non-profits and industrial labs are far more prevalent in full democracies. Authoritarian states are more likely to be populated by universities and Academies of Science.

Table 3. Institutions by sector and the EIU Democracy index.

Sector	Full Democracy	Flawed Democracy	Authoritarian
# Institutions	N=1814	N=1786	N=872
Hospital	16.2%	13.5%	2.6%
Non-profit	6.2%	1.5%	0.6%
Industry	10.6%	4.6%	2.4%
Government	17.3%	18.5%	9.3%
Academia	49.7%	61.3%	72.7%
Academy of Science	0.1%	0.6%	12.4%

As noted earlier, all types of institutions except hospitals tend to have missions that are more well aligned with document-based structures than journal-based structures. The tendency to pursue document-focused missions is not dependent on national context once one takes this into account. National context, especially the tendency for the nation to be more democratic, is mostly influencing whether hospitals are active in publishing in the scientific and technical literature. In autocratic nations, there are far fewer hospitals and non-profit organizations that publish scientific and technical articles.

DISCUSSION

In order to properly compare our results to those from May (1997), King (2004) and others, we will need to cluster and label the 114 DC2 document categories. We will also need to validate that organizations in these groups of document clusters are following the same general mission. This is a work in process. Our preliminary results are encouraging. There are groups of DC2 document categories that correspond to the old industrial revolution and others that correspond to more modern medical practice. Overall, our initial impression is that these groups of DC2 document categories are not disciplinary in the traditional sense. Rather, groups of DC2 document categories represent the broader high-level societal problems that interdisciplinary groups of scientists are trying to solve.

We are interested in understanding why institutions and nations pursue different research agendas. To date, we have not found any literature that looks at publication profiles of institutions as a way of detecting what the missions of these institutions might be. Nor have we found empirical investigations into how these institutional missions might differ according to economic and political contexts. We intend to continue work on this research question and invite others to join us in this effort.

REFERENCES

- Doré, J.-C., Ojasoo, T., Okubo, Y., Durand, T., Dudognon, G., & Miquel, J.-F. (1996). Correspondence factor analysis of the publication patterns of 48 countries over the period 1981-1992. *Journal of the American Society for Information Science*, 47(8), 588-602.
- King, D. A. (2004). The scientific impact of nations. *Nature*, 430, 311.
- Klavans, R., & Boyack, K. W. (2016). Which type of citation analysis generates the most accurate taxonomy of scientific and technical knowledge? *Journal of the Association for Information Science and Technology*, forthcoming.
- May, R. M. (1997). The scientific wealth of nations. *Science*, 275, 793-796.
- Moya-Anegón, F., & Herrero-Solana, V. (2013). Worldwide topology of the scientific subject profile: A macro approach in the country level. *PLOS One*, 8, e83222.



Science policy through stimulating scholarly output. Reanalyzing the Australian case*

Peter van den Besselaar[#], Ulf Heyman[#], Ulf Sandström^{###}

[#] p.a.a.vanden.besselaar@vu.nl
Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

^{##} ulf.heyman@uadm.uu.se
Uppsala University, Uppsala, Sweden

^{###} ulf.sandstrom@indek.kth.se
KTH Royal Institute of Technology, Stockholm, Sweden

Abstract

There is a long standing debate about perverse effects of performance indicators. A main target is science policy using stimulation of output as instrument. The criticism is to a large extent based on a study of the Australian science policy in the early 1990s. Linda Butler studied the effects and argued that the effect was an growth of output, but also a decrease of average quality of the output. These results have been cited many times. In this paper we reanalyze this case and show that the analysis of Butler was wrong: the new Australian science policy did not only increase the output of the system, but also the quality went up. We discuss the implications.

Introduction

More than ten years ago, Linda Butler (2003a) published a well-cited article claiming that the Australian science policy in the early 1990s made a mistake by starting to award the *number* of publications. According to Butler, the policy stimulated researchers to publish more but less good papers. To illustrate this, she first of all showed that Australian number of papers was increasing, as was the share in world production of papers, but that the relative citation impact of those publications did not increase. As the same indicators of other countries tended to increase, Butler concludes that the Australian knowledge production was losing quality. The explanation of this finding is rooted in two behavioral mechanisms that were suggested by Butler (2003b, 2005). Firstly, the increase of Australian publications in top journals (defined in terms of the impact factor) was much slower than the increase of Australian papers in lower impact journals. This suggested that the new policy was stimulating Australian researchers to select on average lower level journals for their increased output. Secondly, the new output oriented policy would stimulate ‘salami slicing’ of papers to increase productivity. Salami slicing stands for publishing results not in one or a few, but in a series of articles, e.g., in a mix of important core journals and peripheral journals, and in journals of a scientist’s own field and in journals belonging to other fields, and in international and national journals. But also for slicing the analysis in different parts, and publishing each in a separate article, instead of all together in a single and longer article (Broad 1981). The lesson was obvious: stimulating quantity of output seems detrimental for the quality of the knowledge production, a lesson

* The authors acknowledge support of the Bank of Sweden Tercentenary Foundation (Searching for excellence and funding effects project; grant P12-1302:1)

that has been cited many times since. And it implies that scientists opportunistically react on *perverse* incentives and try to maximize their indicator scores – instead of maximizing the quality of their research, which is in contrast with many studies showing high motivation and commitment of academic researchers with their work.

From a policy perspective one would want to stimulate the production of (highly cited) papers that make a difference with new ideas, designs, or methods. But how would that work? Highly cited papers are sparse. Simonton (2004) argues that there is randomness in the process, so on average, the more tries, the more successful tries. At the individual level, we found exactly this: the more papers, the more highly cited papers (Van den Besselaar & Sandström 2015; Sandström & Van den Besselaar forthcoming). If it would also work in this way at the aggregate level, the results of Butler would be unexpected. Therefore, the aim of the paper is to re-analyze the effects of the Australian policy intervention in the early 1990s, when a performance (output) based funding system was introduced. This is especially important, as Butler's argument has been repeated many times (e.g., Geuna & Martin 2003; Hicks 2009; OECD 2010; Hicks et al, 2015; cf. Stephan 2012), and has become a part of the received knowledge in science policy studies.

In the remainder of the paper, we will replicate Butlers' study. Using comparable but not identical data, some additional indicators, and a longer time perspective, we investigate whether the 'quantity policy' of Australian government indeed resulted in a decline of the quality of Australian science. We use more up to date quality indicators and a longer time period to investigate whether the pattern we find is indicator-independent. Finally, we address the behavioral aspects. Before we do this, we first reconstruct the analysis of Butler *with our data and indicators*, to show that different outcomes are not due to the differences in data and indicators. This is no secondary analysis of *her* data, but a study of the same phenomenon (the effect of the Australian quantity driven policy on the quality of research) using the same database (Web of Science WoS) and in our view relevant quality indicators. In the next section the differences and similarities with Butler's study are discussed.

Data and Methods

We use WoS-data from InCites. These data are based on full counting - but it is the same method as Butler used.¹ InCites deploys an open citation window and this will, because of the much longer citation window in our data, create some small differences compared to Butlers data. Our analysis covers the same publication types as was used in the Australian funding scheme. We compare the development of Australia with the same set of developed countries that Butler used as frame of reference: Belgium, Canada, England, France, Germany, Italy, Japan, Netherlands, Sweden and Switzerland. Finally, we use for the reconstruction of Butlers' analysis the same time period. But as we have a longer time series at our disposal, we will use that too for the evaluation of the Australian reform announced in 1993. We use the Relative Citation Impact (RCI) – as did Butler. But we also use the percentage of Australian papers within the (field normalized) top 10% most cited papers (PP-top10%), and the field normalized citation score (NCS).

At the behavioral level, we do two brief tests.² Firstly, we distinguish between merely *Australian journals* in the WoS, and *international journals*.³ This is also done using InCites. We then can answer the question whether the Australian growth merely comes from the first type of journals, as suggested by Butler (2003b) (the less good and lower impact journals). We also do a brief analysis of the salami strategy. If Australian researchers would have used this strategy, we would expect that some years after the implementation of the new 'quantity policy' a downward trend can be observed in the Australian paper length.⁴

Publications in Australia 1981-1999: reconstructing Butler’s analysis

Butler’s analysis was largely based on SCI, while our data includes the larger part of WoS (including the Social Science Citation Index and the AHCI), which may introduce some differences in the results. Differences are also expected due to a much wider citation window and due to changes in the database. And we do not fractionalize citations, which at least partly explain the higher citation level in our Figure 1b. In our data there is a decrease in the share of publications from 1987 to 1991, which is not visible in Butler’s figure but the decrease disappears if the comparison is made to similar countries rather than to all countries (this is discussed further below). The difference between the time series may partly be a result of the inclusion of non-western journals into WoS, which would increase the world total of publication and, although to a lesser extent, citations. Comparing Figures 1a and 1b, the share of publications is somewhat lower and the share of citations is higher in our Figure 1b, but the trends look quite the same.

Butler calculated the Australian relative citation impact RCI “by dividing Australia’s share of world publications by its share of world publications”. From the Figure 1a below it is very difficult to observe the yearly variation, since the scale of the Y-axis is unsuitable for an indicator that varies as little as RCI and as similar indicators do. Therefore, her graph easily becomes misleading, as Figure 1c shows. But even using the scale available the trend in Butler’s RCI data is similar to ours (again comparing Fig 1a and Fig 1b).

Figure 1a: Butler’s (2003) figure 1

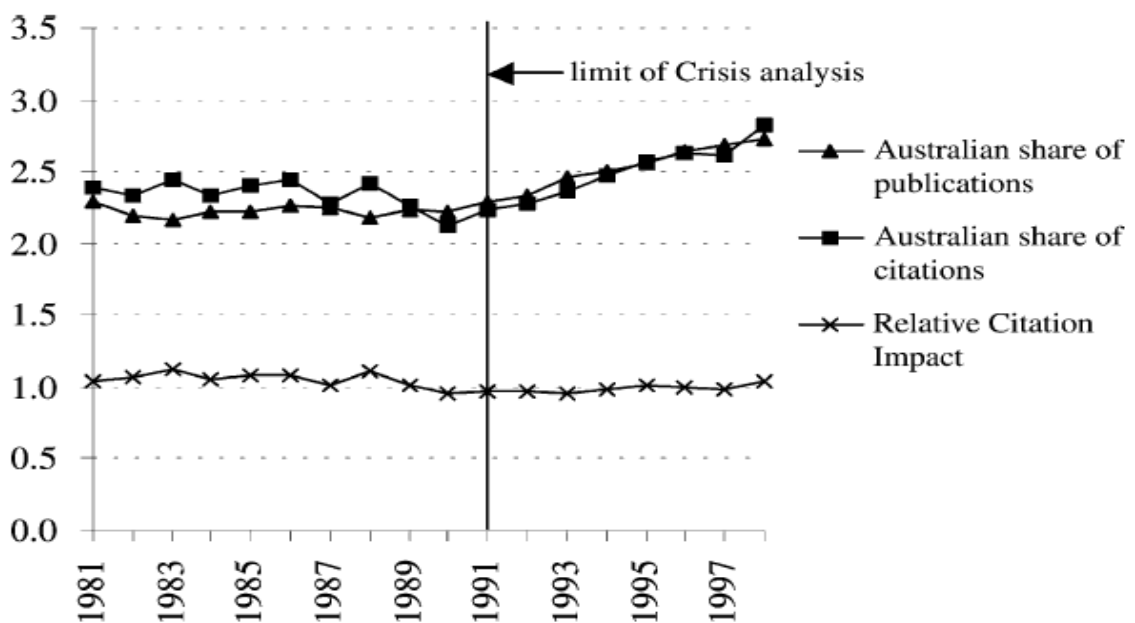
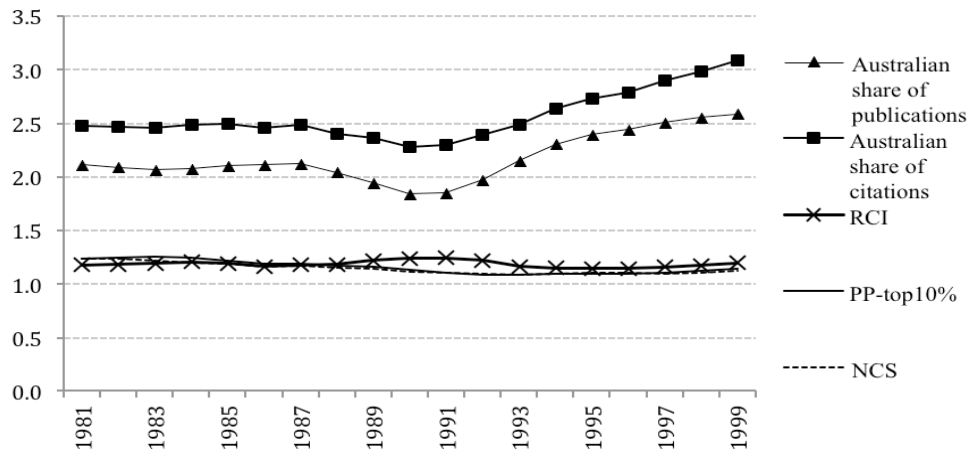


Fig. 1. Australia’s share of publications and citations, and RCI in the SCI, 1981–1998.

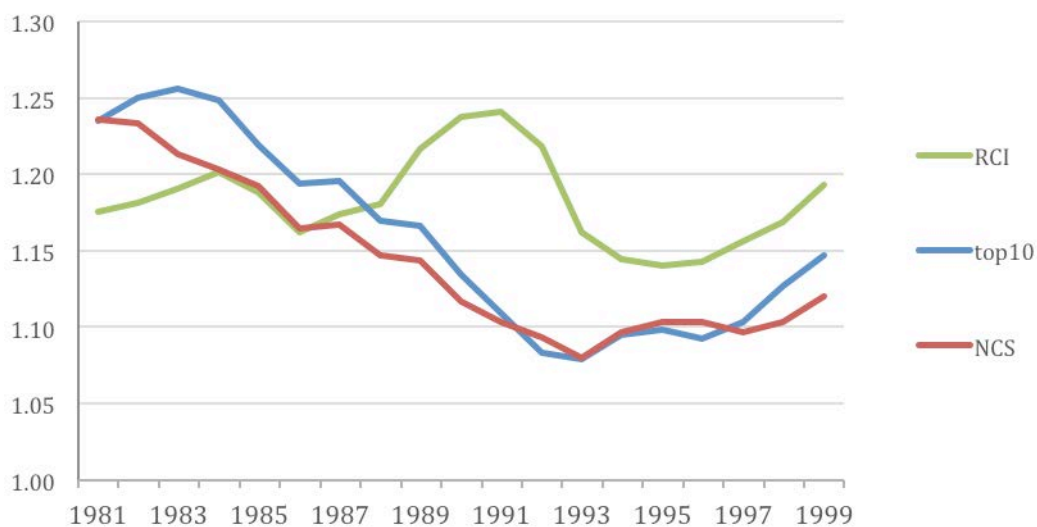
Figure 1b. Our reconstruction of Butlers figure 1



There are also a few differences. Our data show a decline of publications *and* citations from 1987 onwards, which seem to confirm a quite sharp decline in Australia’s competitiveness, while Butler’s data only show a decline in citations. As to the (field and publication type normalized) relative impact (NCS; PP-top10%) our data clearly show a continuous decrease during the 1980s, followed by a period of stabilization and a *tendency to increase* from 1996 onwards (Figure 1b, and more clearly with an appropriate scale in Figure 1c). The RCI we collected from InCites does fluctuate⁵, but as the other two impact indicators it shows an increase starting in 1996, the moment when the new policy was implemented. So the changes in the research system seem *to have had a positive effect* on the impact of Australian publications, according to our indicators.

In order to achieve a better understanding of the effects of the changes in the funding system, we will proceed with a more detailed analysis of the variation of the relative impact, and deploy longer time series.

Figure 1c. Impact of Australian papers, 1981-1999 (several indicators, appropriate scale)

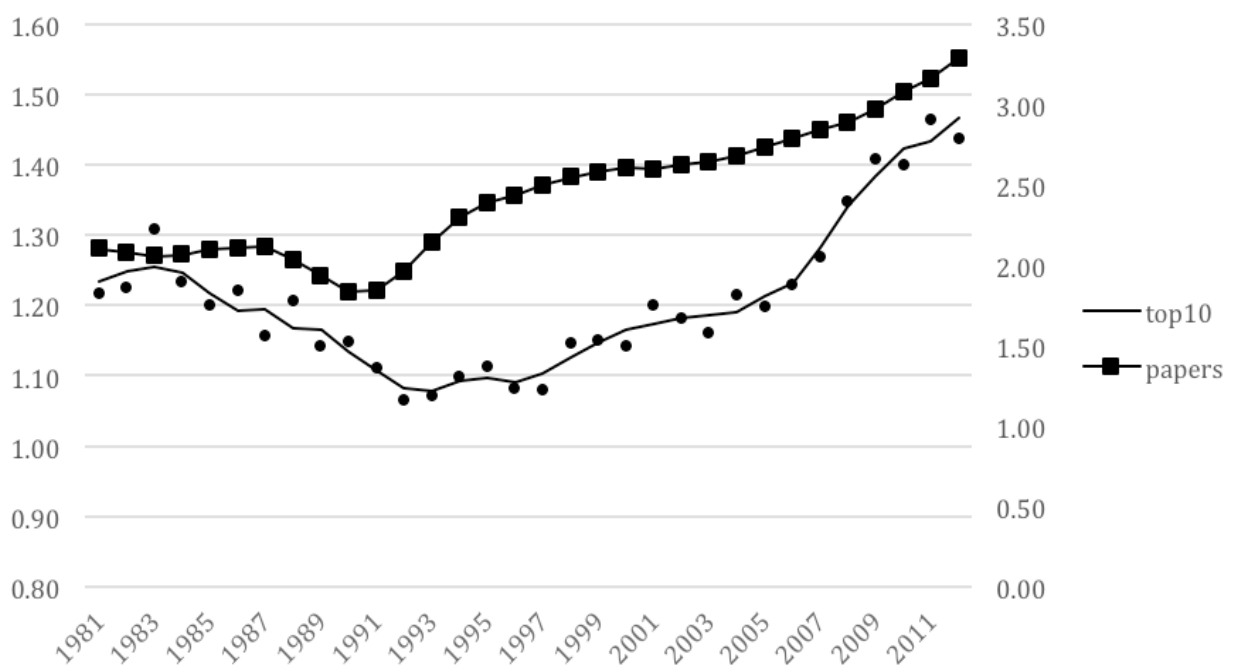


Extending the analysis

A comparison of the number of papers with at least one Australian address to the total number of paper in the world (measured from WoS) shows an impressive increase starting with the announcement of the new science policy in 1993 (Figure 2). The relative share of top 10% cited papers increases (with a delay) even faster, and there is no evidence of decreased quality as a result of changes in the research system. Using approximately the same method as Butler, but extending the time series thus refutes her statement that quality is sacrificed for quantity; on the contrary the indicator for quality increases even more than quantity.

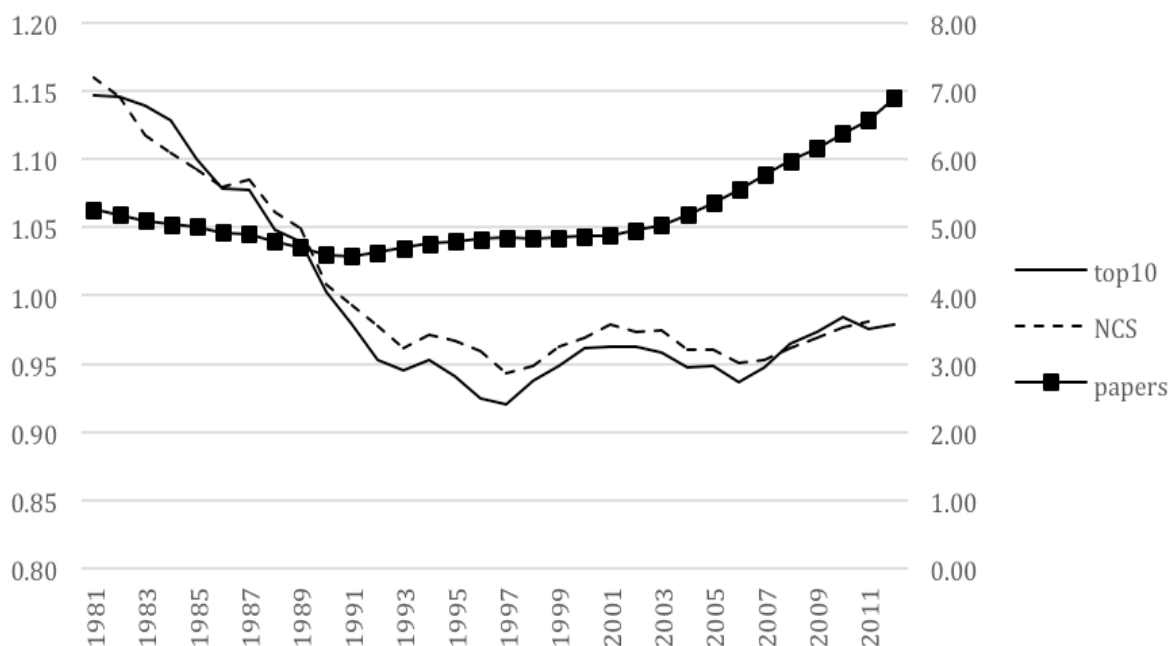
What does this indicate? Obviously, the Australian production was about in par with the world increase of papers. For a developed country that level of performance was not a sign of strength, and the share of top10% cited articles decreased continuously since the early 1980s. There was obviously a good reason for a new science policy, and (in line with the analysis of Butler) we see almost immediately an increase of research output, faster than the world production (Figure 2). In contrast to Butler, we also find that the selected *quality* parameter stopped to decrease in 1993 and started to recover shortly after. The new ‘quantity policy’ that was announced in Australia 1993 and introduced 1995/6 thus seemed to have a significant positive effect on quality. The share of top10 % cited papers with Australian authors increases too at a rather fast pace – suggesting that there is a strong connection (and not a tension) between quality and quantity of the scholarly output.

Figure 2: Growth of the Australian research output and impact (compared to the world average)



Top10% cited papers = left axis; Papers = right axis; 3 years moving average.

Figure 3. Australia's research output and impact (compared to the ten reference countries)



Top10% cited papers, NCS = left axis; Papers = right axis; 3 years moving average.

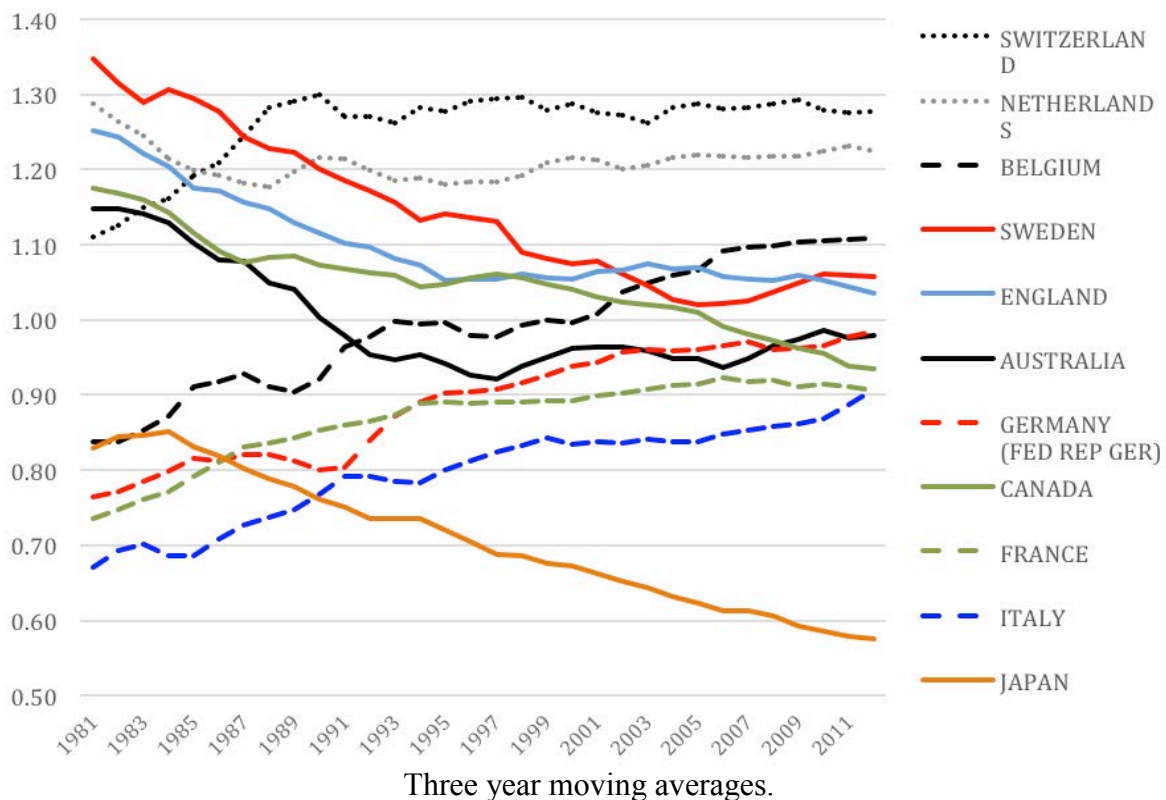
The comparison to world production of scientific papers is however seriously flawed because of increased international cooperation, strong development of scientific publication in emerging countries and rather large changes in the database by inclusion of many new journals, proceedings and later on also books and chapters in books. In order to minimize the effect of these errors, we also compare Australia with a set of similar countries; the same set as Butler used: Belgium, Canada, England, France, Germany, Italy, Japan, Netherlands, Sweden and Switzerland. Compared to this set of reference countries, the development of Australian publication is still positive (Figure 3).

So, again the data indicate that the changes in the science policy did stop the decline of research quality and turned it into a quality increase, not unexpected with some time lag. Summarizing, the recovery of Australia is strong compared to world, and modest but still positive compared to the average of selected developed countries.

Did Australian science develop differently from other countries?

Studying the development of the individual countries yields a rather clear picture (Figure 4). In the beginning of the eighties, the countries were divided in one highly cited group (Sweden, Netherlands, England, Canada, Australia and Switzerland) and one less cited (Belgium, France, Germany Italy and Japan). Since then Netherlands and Switzerland have kept their high share in the top 10% highest cited publications, whereas the four other countries show a decline. Out of these four only Australia and possibly Sweden have reversed the trend and seem to stabilize.

Figure 4. The relative share of top 10% highly cited papers compared to the average of the selected countries.

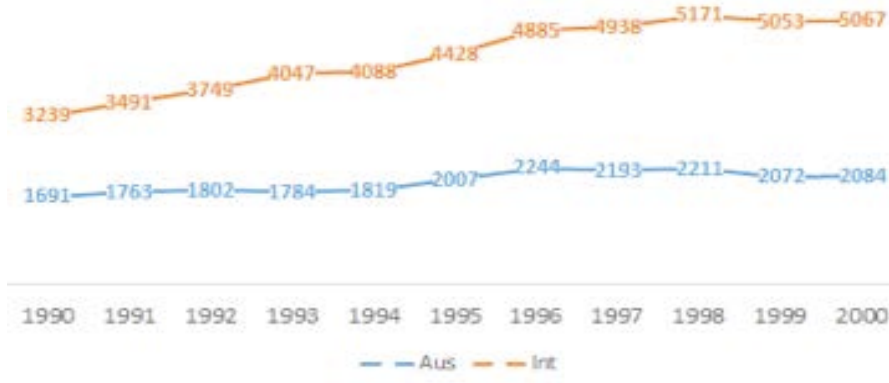


In the low cited group all countries except Japan have increased their share in the top 10% highly cited publications and thereby of course caused Australia's relative decline. The picture suggests a substantial convergence in the world science system, with some exceptions at the higher end (Netherlands, Switzerland) and at the lower end (Japan). Seen in this context, Australia after the 1995-change in the funding system is doing relatively well, and keeps standards at the international average.

Behavioral patterns.

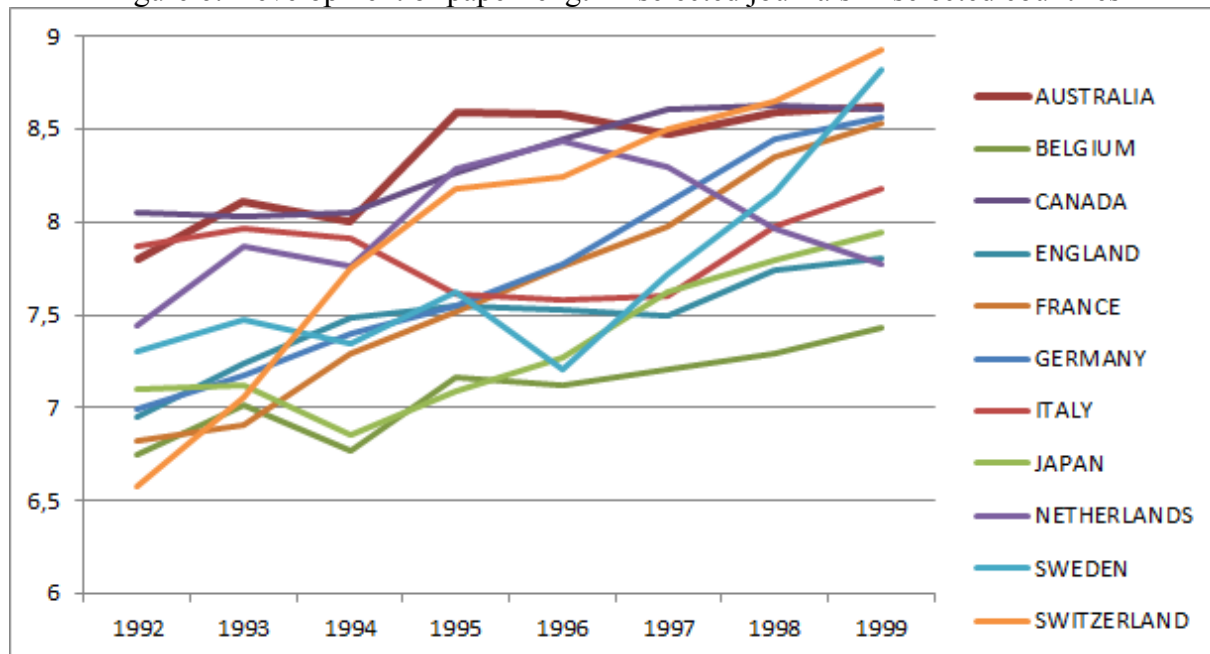
Did Australian researchers move to low level (often more local) journals? Figure 5 shows the growth of Australian output, split into international and local journals – all included in the WoS. Obviously, the growth is mainly in international journals, indicating a reorientation of Australian science towards internationalization.

Figure 5: Australian output in international and local journals



The second behavioral pattern attributed to ‘quantity-policies’ is ‘salami slicing’ of papers: reducing papers to the smallest publishable unit. Did Australian researchers start to do this, in order to increase output levels? One would expect that salami slicing would lead to at least some reduction of paper length (Broad 1981). We calculated the (weighted) average length in three journals (*Plant and Soil*, *Astrophysical Journal* and *Biochemical Journal*). We took a specific set of journals in order to control for composition effects, as paper length differs by field. The first results indicate no trend towards salami slicing: the Australian papers do not show a downward trend after the new policy was implemented, and they are actually relatively long compared to the selected set of other countries (Figure 6).

Figure 6: Development of paper length – selected journals in selected countries



Conclusions

Australia shows a drop in performance in the beginning of the nineties, for publication share and also for citation based quality indicators: the RCI and the share top10% cited papers. It had been a continuous trend since beginning of the 1980s, and all parameters show a recovery from around 1993 – with as expected citation indicators following with some delay. The

recovery is especially strong compared the whole world, and weak but positive compared to the selected set of scientifically strong countries.

When comparing individual countries, the decline since the beginning of the 1980s does not seem to be a specific problem for Australia, as we also observe this for Sweden, England and Canada. It seems an effect of the convergence of research performance of developed countries, with only a few countries that escape from this – in positive or negative direction. Furthermore, data suggest that the switch in publishing strategies in Australia was not focusing on the local and regional journals, but towards international ones. And we do not find indicators that the Australian academics started to salami slice papers, in order to increase output.

What do we learn from this? Of course, performance based evaluation is not the only factor influencing output and quality levels; other factors such as funding levels, and the organization of research and of the higher education system may be important too (Sandström & Van den Besselaar, in preparation). Nevertheless, the data suggest that the new policy in the early 1990s gave the system a new impulse – as funding became output-dependent. This initiative did not only result in higher productivity, but obviously also in higher quality. Using a longer time series and scales that match the variation of the variables it becomes obvious that there is a positive and not a negative relationship between the size of the output and its quality. Quantity matters, not only at the level of individual researchers (Van den Besselaar & Sandström 2015; Sandström & Van den Besselaar, forthcoming; Larivière & Costas 2015) but also at the level of the science system. If researchers are stimulated to become more productive, to produce more ideas (and papers), they also produce more good ideas (and more good papers). This is an important finding - also for policy. And it stands in strong contrasts to calls for less focus on publications (The Leiden Manifesto). Much of the critique on performance based funding seems wrong and at least strongly overstated – often uncritically based on the false conclusions about the Australian experiences (e.g., Hicks 2009; Hicks et al 2015, OECD 2010).

References

- Broad, W.J. (1981). "The Publishing Game: Getting More for Less", *Science* **211** (13 March), 1137–1139, doi:10.1126/science.7008199
- Butler, L. (2003a). Explaining Australia's increased share of ISI publications—the effects of a funding formula based on publication counts. *Research Policy* **32** (1), 143–155.
- Butler L. (2003b). Modifying publication practices in response to funding formulas. *Research Evaluation* **12** (1), 39-46.
- Butler, L. (2004). What happens when funding is linked to publication counts? In H. F. Moed, W. Glanzel, & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology* (pp. 340–389). London: Kluwer Academic Publishers.
- Butler L. (2010). Impacts of performance-based research funding systems: A review of the concerns and the evidence. In OECD (2010). *Performance-based Funding for Public Research in Tertiary Education Institutions: Workshop Proceedings*, 127-165.
- Geuna A., Martin B (2003). University Research Evaluation and Funding: An International Comparison, *Minerva* **41**, 277-304.
- Hicks D. (2009). Evolving regimes of multi-university research evaluation. *Higher Education* **57**, 393-404
- Hicks D, Wouters P, Waltman L, de Rijcke S & Rafols I (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, April 22.

- Larivière, V., Costas, R. (2015). How many is too many? On the relationship between output and impact in research. In: *Proceedings of the 15th International Conference of the International Society for Scientometrics and Informetrics*, pp. 590-595.
- Sandström U & Van den Besselaar P, On Quantity and Quality; The importance of publishing many papers (*forthcoming*)
- Sandström U & Van den Besselaar P, Funding modes and the performance of national research systems (*in preparation*)
- Simonton, Dean Keith (2004). *Creativity in Science: Chance, Logic, Genius, and Zeitgeist*. N.Y. Cambridge Univ Press. [Reprinted 2008].
- Stephan, P (2012). Perverse incentives. *Nature* **484**, April 5, pp.29-31
- OECD (2010). *Performance-based Funding for Public Research in Tertiary Education Institutions: Workshop Proceedings*.
- Van den Besselaar P, Sandström U (2015). Does quantity make a difference? *Proc. ISSI 2015 conference* (ed by A. Salah, S Sugimoto, U. Al), Istanbul, 577-583

¹ This introduces a number of difficulties when countries are compared. We partly address these below.

² In a next paper, we plan to extend this part of the analysis, and investigate salami slicing in a larger set of international journals.

³ Australian journals are defined as those journals with a reference to Australia or synonym in the journal title, e.g. Australian Journal of Political Science. There are 55 such journals in the Web of Science over the period. International journals are, in this case, all other journals in WoS.

⁴ A few years after, as it takes a while for this becomes visible: papers need firstly to be written, and then it takes time for review and publishing.

⁵ Butler's RCI (Figure 1a) has a maximum in 1991. The difference with our Figure 1b can be understood, as we use field-normalized and field-weighted figures. The Butler RCI is an unweighted indicator that to a large degree depended on the large biomedical fields - as the indicator was built on share of citations divided to share of publications.



The need for contextualized scientometric analysis: An opinion paper¹

Ludo Waltman and Nees Jan van Eck

waltmanlr@cwts.leidenuniv.nl; ecknjpvan@cwts.leidenuniv.nl
Centre for Science and Technology Studies, Leiden University, Leiden (The Netherlands)

ABSTRACT

INTRODUCTION

Scientometric indicators, in particular indicators based on citations, nowadays play a prominent role in research evaluations. Given the importance of these indicators, scientometricians are putting a lot of effort into making technical improvements to the indicators in order to increase their accuracy. Especially indicators based on citations have received a lot of attention during recent years. This has resulted in the development of many advanced citation-based indicators (for a review of the literature, see Waltman, 2016). At the same time, scientometricians have been exploring all kinds of new indicators, many of which are referred to as altmetric indicators. Interest in these new indicators is largely driven by the availability of new data sources, but also seems to relate to changing viewpoints on research evaluation, in particular an increasing focus on evaluating the societal impact of research. Like in the case of traditional citation-based indicators, scientometricians are trying to obtain more and more accurate statistics by developing increasingly advanced indicators (e.g., Fairclough & Thelwall, 2015; Haunschild & Bornmann, 2016).

The scientometric community also regularly discusses the role that scientometric indicators should play in research evaluations, for instance in evaluations of the scientific performance of research groups or research institutions. Although there is no full consensus on this issue (e.g., Abramo & D'Angelo, 2011), most scientometricians seem to agree that peer review should be the primary instrument for research evaluation, while scientometric indicators can be used to support peer review (e.g., Butler, 2007; Moed, 2007; Van Raan, 2004). This approach is referred to as informed peer review. Recent statements in favor of informed peer review were for instance made in the Leiden Manifesto (Hicks et al., 2015), which presents ten principles for the use of quantitative indicators in research evaluations, and in a report on the role of quantitative indicators in the Research Excellence Framework in the United Kingdom (Wilsdon et al., 2015).

We have now sketched two important lines of thinking in the scientometric community. On the one hand, there is the focus on indicator development, where scientometricians develop both completely new indicators and more advanced variants of existing indicators. On the other hand, there is the idea of informed peer review, according to which indicators should be used to support peer review rather than to replace it. In the first part of this opinion paper, we

¹ We would like to thank our colleagues at the Centre for Science and Technology Studies (CWTS) of Leiden University for useful discussions that have contributed to the development of the ideas presented in this paper. In particular, we are grateful to Paul Wouters for his suggestions that have had a significant influence on our ideas.

reflect on these two lines of thinking. We start by discussing the traditional approach to informed peer review, which relies strongly on the use of advanced citation-based indicators. We argue that this approach fails to combine expert judgment and scientometric indicators in a truly integrated way. As a consequence, there is a need for an alternative approach to informed peer review. In the second part of the paper, we propose such an alternative approach. Our proposed approach, which is based on the principles of context, simplicity, and diversity, is referred to as contextualized scientometric analysis.

We note that in this opinion paper we present our personal ideas on the use of scientometric indicators in research evaluations. We do not aim to offer systematic scientific evidence to support the ideas that we present. The observations that we make in this paper are based on a number of years of experience that we have both with the scientific study of scientometric indicators and with the practical application of these indicators.

PROBLEMS OF THE TRADITIONAL APPROACH TO INFORMED PEER REVIEW

Scientometric indicators can provide useful information to support peer review. Consider a peer review committee that is responsible for evaluating the scientific performance of a research group. Most likely, it will not be feasible for the committee to evaluate in full detail all scientific outputs of the research group, and in fact the committee may not even have sufficient expertise for such a comprehensive evaluation. In-depth peer review will therefore be restricted to a selection of the scientific outputs of the research group. Scientometric indicators can then be used by the peer review committee to complement the results of in-depth peer review with quantitative information, especially for scientific outputs that have not been evaluated in detail by the committee. Furthermore, because the peer review committee will include only a limited number of experts, its evaluation of the scientific performance of the research group may not be representative of the assessment of the group's performance by the scientific community at large. Scientometric indicators can be used by the peer review committee to get an impression of the broader reception of the work of the research group.

Although information provided by scientometric indicators can be helpful to support peer review, the use of these indicators to replace peer review would in general be problematic. This is because scientometric indicators have major limitations. Some dimensions of scientific performance can be quantified reasonably well using scientometric indicators. The typical example is the scientific impact of publications, which can be quantified using citation-based indicators. However, other dimensions of scientific performance, such as productivity (Waltman et al., in press) and societal impact (Bornmann, 2013), are much more difficult to quantify using scientometric indicators. In addition, indicators that work reasonably well in some fields of science may not work well in others. For instance, citation-based indicators work quite well in most of the sciences, but their value in the social sciences and humanities is more questionable. Given the limitations of scientometric indicators, these indicators generally should not be used to replace peer review. For this reason, most scientometricians recommend the idea of informed peer review, in which scientometric indicators are used to support peer review.

Scientometricians are less clear about the specific way in which informed peer review should be implemented. In the simplest approach, which seems to be the one that is most commonly used in practice, in the beginning of a peer review process, the experts carrying out the peer review receive a report presenting selected scientometric indicators for the research unit to be evaluated. The experts then decide themselves how the information provided by the indicators

is used in the peer review process. The integration of indicators in the peer review process tends to be limited in this approach. For instance, when new questions arise during the peer review process, there usually is little or no opportunity for the experts to ask for additional scientometric information.

When this traditional approach to informed peer review is taken, scientometricians typically recommend the use of advanced scientometric indicators, for instance field-normalized citation-based indicators calculated based on Web of Science or Scopus data. They usually warn against the use of simple citation-based indicators such as the h-index (Hirsch, 2005) and the journal impact factor. Scientometricians also tend to be critical about the use of indicators based on Google Scholar and altmetric indicators. The usefulness of these indicators is questioned because they rely on low quality data sources, because they can be manipulated relatively easily, and in the case of altmetric indicators also because it is not sufficiently clear what they measure (e.g., Thelwall & Kousha, 2015; Wouters & Costas, 2012). In practice, informed peer review therefore usually means that experts are provided with a limited set of scientometric indicators, mainly advanced citation-based indicators that aim to quantify the scientific impact of publications.

We believe that this traditional approach to informed peer review has two important problems. The first problem is that scientometric indicators are provided only for specific dimensions of the scientific performance of a research unit, primarily the scientific impact dimension. Relevant scientometric information on other dimensions often is not used, because indicators constructed based on such information, for instance altmetric indicators, are not considered to be of sufficient quality. This means that relevant information is ignored, even though this information is available. It also creates the risk that experts may put too much emphasis on a single dimension of scientific performance, not because they regard this as the most important dimension, but because this is the dimension for which they have scientometric indicators available.

The second problem relates to the focus of scientometricians on advanced indicators, in particular field-normalized citation-based indicators. Advanced indicators can be expected to be more accurate than simple indicators, but this increase in accuracy comes at a cost. To illustrate this, let us take field-normalized citation-based indicators as an example. Compared with non-normalized indicators, these indicators are more accurate because of the correction they make for differences in citation practices between fields. However, this does not mean that field-normalized indicators have a perfect accuracy and that experts in a peer review committee can blindly trust these indicators. Instead, in order to assess the accuracy of the indicators, experts need to have a good understanding of the way the field normalization has been carried out. For instance, when field normalization is carried out based on predefined fields in the Web of Science or Scopus database, experts need to be aware of the inaccuracies of these fields and of the biases these fields may cause (Leydesdorff & Bornmann, 2016; Van Eck et al., 2013; Wang & Waltman, 2016). This of course means that we ask a lot from experts. Scientometric indicators aim to make life easier for experts by providing summarized information that hides some of the complexity of the real world. However, at the same time, these indicators, especially the more advanced ones, introduce new complexity that experts need to deal with. This new complexity consists of the knowledge experts need to have to properly interpret scientometric indicators and to understand their limitations.

The use of advanced scientometric indicators is challenging also because it is difficult for experts to directly link these indicators to the concrete real-world entities on which their expert judgment is based. For instance, field-normalized citation-based indicators provide statistics that result from quite complex calculations. Although the calculations start from concrete real-world entities, namely the publications and citations of a research unit, the statistics obtained at the end of the calculations lack a direct connection to these concrete real-world entities. Field-normalized indicators therefore have a high level of abstractness. As a consequence, experts usually interpret these indicators not by relating them to their own expert judgment, but by following simple guidelines suggested by scientometricians.

As an example, consider the mean normalized citation score (MNCS) indicator (Waltman et al., 2011), which is the average field-normalized number of citations of the publications of a research unit. The MNCS of all publications worldwide in a field of science equals one. A simple guideline therefore is that, in terms of scientific impact, a research unit with an MNCS above one performs above world average, or above expectation, while a research unit with an MNCS below one performs below world average, or below expectation. On the one hand, such a simple guideline demonstrates the power of advanced scientometric indicators. The complexity of the real world is reduced into a single number with an easy-to-understand interpretation. On the other hand, we also believe that such a simple guideline may have undesirable effects. It may draw attention away from the details of the calculation of an indicator, reducing the awareness of the limitations of the indicator. It may also discourage experts to reflect more in depth on the interpretation of an indicator. For instance, in the interpretation of the MNCS indicator, experts may blindly follow the above-mentioned guideline without thinking in any detail about the specific publications of a research unit that have been cited a lot and the reasons why these publications may have been cited so many times.

Because of the problems discussed above, we believe that the traditional approach to informed peer review fails to combine expert judgment and scientometric indicators in a truly integrated way. In the next section, we propose an alternative approach to informed peer review. This approach, referred to as contextualized scientometric analysis, aims to achieve a better integration of expert judgment and scientometric indicators.

AN ALTERNATIVE APPROACH TO INFORMED PEER REVIEW: CONTEXTUALIZED SCIENTOMETRIC ANALYSIS

Contextualized scientometric analysis is based on three fundamental principles: Context, simplicity, and diversity. We start by discussing each of these principles:

- *Context.* Scientometric indicators should be complemented with contextual information. The scientometric context of an indicator consists of all scientometric information that can be relevant in the interpretation of the indicator. When indicators are made available, their scientometric context should also be made available as much as possible.

For instance, suppose we report the average number of citations of the publications of a research unit. Relevant information to support the interpretation of this indicator includes the list of publications of the research unit and each publication's number of citations. Especially the most highly cited publications are important, because these publications largely determine the value of the indicator. Therefore a list of the most highly cited publications of the research unit should be made easily accessible. Experts

then can use this list to better interpret the indicator. They may for instance use the list to find out whether the most highly cited publications of the research unit are research articles or review articles, whether they have been co-authored with external collaborators, and in the case of co-authorship, whether the research unit has been one of the leading authors. If experts find out that a research unit has received most of its citations from review articles, the experts may not be very impressed by the fact that the research unit has a large average number of citations per publication. On the other hand, if the most highly cited publications of a research unit all turn out to present original research that has been carried out without the help of external collaborators, a large average number of citations per publication may be seen as a significant accomplishment.

As a second example of the value of contextual information, suppose we work with a field-normalized indicator. To properly interpret such an indicator, it should be clear how fields are defined. The indicator should therefore be complemented with information on the definitions of fields. For each publication of a research unit, this information could indicate the field to which the publication is assigned. Also, for each field, the main research topics covered by the field could be shown, possibly in a visual way. When experts are aware of the field definitions used by a field-normalized indicator, they can give a more meaningful interpretation to the indicator.

- *Simplicity.* Scientometric indicators should preferably be simple and easy to understand. As we have discussed, advanced indicators introduce complexity and abstractness, which is problematic from the point of view of the integration of indicators in peer review processes. For this reason, whenever possible, simple indicators should be used instead of more advanced ones. For instance, when an analysis is restricted to a single field of science, there is no clear need to use advanced field-normalized indicators and therefore it may be better to use simple non-normalized indicators. Even when an analysis involves multiple fields of science, simple non-normalized indicators may be preferable over advanced field-normalized indicators if the gain in accuracy obtained from field normalization does not offset the difficulties caused by the complexity and abstractness of field-normalized indicators.
- *Diversity.* A diversity of scientometric indicators is needed to cover the different dimensions of scientific performance. Traditional scientometric analyses tend to focus almost exclusively on advanced citation-based indicators calculated based on Web of Science or Scopus data. The problem of this narrow focus is that for various important dimensions of scientific performance no indicators are provided. This is unsatisfactory. Therefore, instead of restricting oneself to a limited set of indicators, a rich diversity of indicators should be made available. This could for instance include indicators based on Google Scholar, altmetric indicators, indicators that take into account input data (e.g., a research unit's number of FTEs or its amount of research funding), indicators based on ratings of journals, conferences, or book publishers, and patent-based indicators. Many of these indicators can be criticized, for instance because of data quality issues, the possibility of manipulation, or the ambiguity of their interpretation. However, when used carefully, these indicators can provide relevant and important information.

The proper use of the indicators is supported by the two principles mentioned above. By complementing the indicators with contextual information and by keeping the indicators as simple as possible, experts are assisted in the interpretation of the

indicators. For instance, based on an altmetric indicator, experts may find out that the publications of a research unit have been mentioned in an exceptionally large number of tweets. The experts could then analyze the context of the altmetric indicator. They may find out that most of the tweets originate from non-scientists interested in the work of the research unit, suggesting that the work may have a significant societal impact. Alternatively, the experts may find out that most of the tweets have been created automatically by bots, and based on this the experts may decide to ignore the altmetric indicator. Hence, by analyzing the context of the altmetric indicator, experts are able to give an appropriate interpretation to the indicator.

We note that the importance of using a diversity of indicators is also emphasized in early work by Martin and Irvine (1983) and more recently in the ACUMEN project (<http://research-acumen.eu>).

In practice, to fully implement the idea of contextualized scientometric analysis, an interactive software tool is needed for making available scientometric indicators and the associated contextual information. Presenting a contextualized scientometric analysis in a traditional static report is hardly feasible. In such a report, it would be difficult to present a large diversity of indicators and it would be even more challenging to make available contextual information to support the interpretation of the indicators. An interactive software tool has the advantage that information can be made available in a selective way based on the specific needs of users. For instance, rather than immediately presenting all available indicators, users can be invited to select the indicators that are most relevant for their particular information needs. These indicators will then be shown, after which users may find out that some of the indicators provide unexpected results that require a more detailed analysis. For these specific indicators, users may then dig into the contextual information to better understand what can be learned from the indicators. For other indicators that turn out to provide less interesting results, users can simply choose not to make use of the available contextual information. So the benefit of an interactive software tool is that a large amount of information can be made accessible in a way that prevents users from being overwhelmed by the volume of the information. In addition, some information can also be presented visually, for instance using visualizations similar to the ones provided by the VOSviewer software (Van Eck & Waltman, 2010).

As already mentioned, contextualized scientometric analysis aims to achieve a better integration of expert judgment and scientometric indicators. This also requires a specific way of working with scientometric indicators. Rather than scientometric indicators playing a more or less independent role in a research evaluation, there should be a continuous interaction between expert judgment and scientometric indicators. For instance, on the one hand, experts may use indicators to help them decide which publications of a research group to read, which researchers to interview, and which questions to ask in the interviews. On the other hand, based on the information experts get from reading publications or interviewing researchers, they may be able to better make sense of the information provided by indicators and they may decide to explore certain indicators in more detail. Ideally, after a process of triangulating expert judgment and scientometric indicators, experts should be able to provide an assessment of a research unit that is supported by and consistent with both their own qualitative judgment and the quantitative information provided by indicators.

CONCLUSION

In discussions on the role of scientometric indicators in research evaluations, most scientometricians advocate the idea of informed peer review. According to this idea, scientometric indicators are used to support peer review but not to replace it. In this opinion paper, we have argued that the traditional approach to informed peer review has two important problems. First, it is focused mainly on providing indicators of the scientific impact of publications and it fails to provide information on other dimensions of scientific performance. Second, it relies strongly on the idea of working with advanced field-normalized indicators. The disadvantage of these advanced indicators is their complexity and abstractness.

We have proposed an alternative approach to informed peer review that aims to achieve a better integration of expert judgment and scientometric indicators. Our proposed approach, which we refer to as contextualized scientometric analysis, is based on three fundamental principles: Context, simplicity, and diversity. According to these principles, a diversity of indicators should be provided, the indicators should be simple, and they should be complemented with contextual information. By following these principles, indicators will better serve the needs of experts and it will be easier for experts to combine the information provided by indicators with their own expert judgment.

When should contextualized scientometric analysis be used? Contextualized scientometric analysis can be expected to have most value in research evaluations in which experts have sufficient time for a detailed analysis of the performance of a research unit. It is also essential that experts recognize the importance of contextual information. Experts should be aware that indicators cannot be blindly trusted, and they should have the curiosity to explore the context of an indicator. When a research evaluation needs to be carried out in a more superficial way and there is less room for a detailed examination of the performance of a research unit, contextualized scientometric analysis has little added value. In that case, a traditional approach to informed peer review, in which a limited set of advanced scientometric indicators are provided, seems preferable. The indicators will probably be interpreted in a more or less mechanistic way, without paying much attention to contextual information. It is then important that the indicators are of the highest possible accuracy, which means that it is best to work with a limited set of advanced indicators.

Our focus in this paper has been on the use of scientometric indicators in research evaluations. However, we believe that the idea of contextualized scientometric analysis may be useful also in applications of scientometric indicators that are not of a strictly evaluative nature. For instance, when a research institution needs to decide on its research priorities for the coming years, it may use scientometric indicators to support its internal decision making process. Again, the best way to benefit from scientometric indicators may be to take an approach that follows the principles of context, simplicity, and diversity.

REFERENCES

Abramo, G., & D'Angelo, C.A. (2011). Evaluating research: From informed peer review to bibliometrics. *Scientometrics*, 87(3), 499–514.

- Bornmann, L. (2013). What is societal impact of research and how can it be assessed? A literature survey. *Journal of the American Society of Information Science and Technology*, 64(2), 217–233.
- Butler, L. (2007). Assessing university research: A plea for a balanced approach. *Science and Public Policy*, 34(8), 565–574.
- Fairclough, R., & Thelwall, M. (2015). National research impact indicators from Mendeley readers. *Journal of Informetrics*, 9(4), 845–859.
- Haunschild, R., & Bornmann, L. (2016). Normalization of Mendeley reader counts for impact assessment. *Journal of Informetrics*, 10(1), 62–73.
- Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature*, 520, 429–431.
- Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46), 16569–16572.
- Leydesdorff, L., & Bornmann, L. (2016). The operationalization of “fields” as WoS subject categories (WCs) in evaluative bibliometrics: The cases of “library and information science” and “science & technology studies”. *Journal of the Association for Information Science and Technology*, 67(3), 707–714.
- Martin, B.R., & Irvine, J. (1983). Assessing basic research: Some partial indicators of scientific progress in radio astronomy. *Research Policy*, 12(2), 61–90.
- Moed, H.F. (2007). The future of research evaluation rests with an intelligent combination of advanced metrics and transparent peer review. *Science and Public Policy*, 34(8), 575–583.
- Thelwall, M., & Kousha, K. (2015). Web indicators for research evaluation, part 2: Social media metrics. *El Profesional de la Información*, 24(5), 607–620.
- Van Eck, N.J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538.
- Van Eck, N.J., Waltman, L., Van Raan, A.F.J., Klautz, R.J.M., & Peul, W.C. (2013). Citation analysis may severely underestimate the impact of clinical research as compared to basic research. *PLOS ONE*, 8(4), e62395.
- Van Raan, A.F.J. (2004). Measuring science: Capita selecta of current main issues. In H.F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of quantitative science and technology research* (pp. 19–50).
- Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10(2), 365–391.
-

Waltman, L., Van Eck, N.J., Van Leeuwen, T.N., Visser, M.S., & Van Raan, A.F. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, 5(1), 37–47.

Waltman, L., Van Eck, N.J., Visser, M., & Wouters, P. (in press). The elephant in the room: The problem of quantifying productivity in evaluative scientometrics. *Journal of Informetrics*.

Wang, Q., & Waltman, L. (2016). Large-scale analysis of the accuracy of the journal classification systems of Web of Science and Scopus. *Journal of Informetrics*, 10(2), 347–364.

Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., ... Johnson, B. (2015). *The metric tide: Report of the independent review of the role of metrics in research assessment and management*. HEFCE.

Wouters, P., & Costas, R. (2012). *Users, narcissism and control: Tracking the impact of scholarly publications in the 21st century*. SURFfoundation.



Fake Academic Degrees as an Indicator for Severe Reputation Crisis in the Scientific Community of Russian Federation

Andrey Rostovtsev*, Alexander Kostinskiy**

*dissernet@gmail.com

IITP RAS, Bolshoy Karetny per. 19, build.1, Moscow 127051 (Russia)

*** kostinsky@gmail.com

National Research University Higher School of Economics, Myasnitskaya str. 20, Moscow, 101000 (Russia)

ABSTRACT

As is shown by our studies described below, awarding fake academic degrees to politicians, businessmen, doctors in clinics, professors in universities, and teachers in schools, that is, to all those who wish to use their new academic titles to step onto a faster career route, is widely accepted in Russia. Fake academic titles are awarded throughout the country. This business is based on the manufacture of falsified dissertations. In early 2013, a group of five researchers and journalists established a social network called “Dissernet”. Dissernet is a volunteer-effort free association aimed at making fraud and trickery in the awarding of academic titles transparent and exposed to the public. The scale of academic fraud in Russia has turned out to exceed the most pessimistic expectations. Statistics collected by Dissernet have yielded a number of conclusions discussed below.

Keywords: plagiarism, Dissernet, Russia, fake academic degree, reputation crisis

DATA AND METHODS

Western societies already have extensive experience in identifying plagiarism through network communities (Weber-Wulff, 2014). VroniPlag, for example, a German project, has already helped identify hundreds of plagiarised dissertations in the last several years. Since the scale of academic fraud in Russia seems to be orders of magnitude larger than the German one, a novel approach has to be developed. Detection of thousands of fraudulent dissertations is mainly the result of a unique technology developed by the Dissernet association. In Russia, along with a thesis (dissertation) a so-called avtoreferat is to be made publicly available before the Ph.D. defense. An avtoreferat consists of the shortened dissertation content (usually 20 to 30 pages) and the main research results. Importantly, the texts of avtoreferats are indexed by public search engines (such as Google or Yandex). The dissertation itself is not usually indexed, however. But if the dissertation contains large fragments of plagiarized text, as described above, its avtoreferat would also have text coinciding with earlier works. Software specially developed for Dissernet is able to pick up avtoreferats one by one and takes advantage of the search engines indices to look for textual coincidences within the whole publicly available corpus of Russian digitized texts, including texts of other avtoreferats. This program runs 24/7. So far, a few hundred thousand dissertations have been automatically checked. Furthermore, Dissernet takes advantage of the common practice of a flow manufacture of fraudulent dissertation. As soon as a rampant plagiarism is detected in a dissertation, it is very likely to be also detected in other dissertations defended in the same dissertation council or with the same supervisor. This happens because the producers of

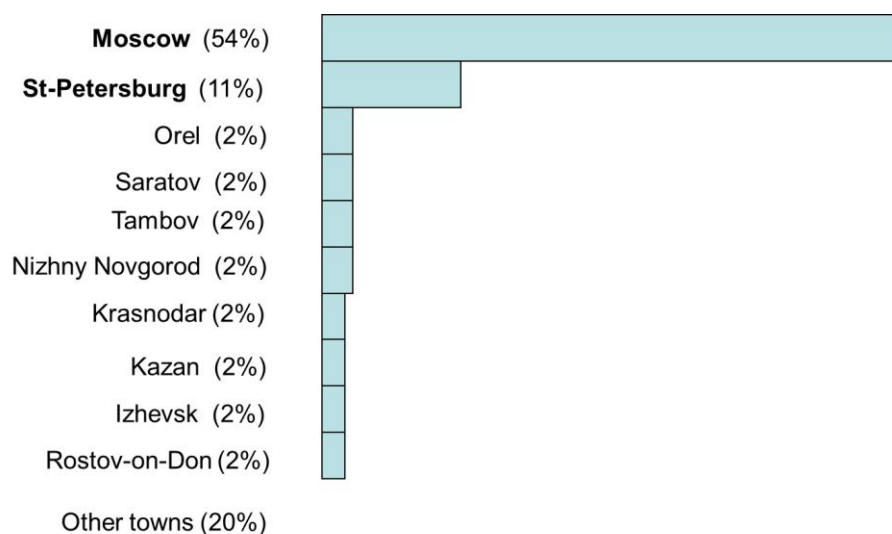
fake dissertations in Russia work in a conveyor-belt mode using very limited sets of scientific texts as sources. In addition, in terms of their involvement into the academic falsification business, it is important to scrutinize dissertation councils' members in those leading universities where PhDs are awarded, the experts in the Higher Attestation Commission, the agency that coordinates and validates the awarding of academic degrees, and last but not least, editorial boards' members in those scientific journals where papers by doctoral degree seekers are to be published prior to the defense. The technique developed by Dissernet allows making such checks automatically. The persons involved directly into awarding fake degrees – more than 10,000 Russian scientists identified so far – are listed in the Dissernet data base (Rostovtsev, 2015).

RESULTS

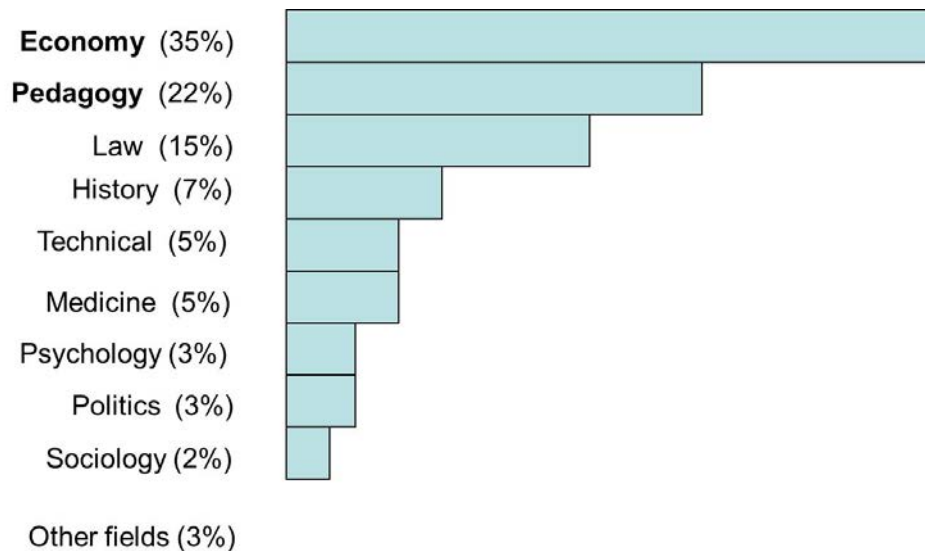
The scale of academic fraud in Russia. Since 2013, Dissernet activists have identified more than 5,250 plagiarized and falsified dissertations defended in the last 15 years. Over 1,120 cases of such dissertations are documented on Dissernet's website (www.dissernet.org).

In terms of geographic location of the universities that award fake degrees, Moscow and Saint Petersburg are playing the most important role as they are major cities, according to the present Dissernet statistics. Other cities and towns fall behind (see Fig. 1).

Figure 1: Geographic locations of universities that award the fake degrees



Distribution across scientific disciplines. According to Dissernet statistics, the number of fake dissertations varies significantly across different academic fields. Most fake dissertations (40%) are defended in economics. Other popular areas are pedagogy and law, followed by medical sciences, political sciences, engineering, and social sciences (see Fig 2). Fake dissertations are rare in the area of natural sciences. Such distribution is symptomatic as it represents Russia's main problem areas: economics, law, education, etc. Moreover, as it is demonstrated, the number of fake dissertations per field is inversely related to Russia's scientific international input in these disciplines.

Figure 2: Fake dissertation distribution across scientific fields.

By focusing on almost totally plagiarized texts, Dissernet only deals with a small tip of scientific fraud in Russia. But even so, in problematic fields such as economics and law, about 3 % of dissertations contain large-scale plagiarism. In pedagogy this fraction is a bit higher, but still below 6 %.

Severe reputation crisis. Last but not least, Dissernet investigations are relevant not only for an assessment of the Russian fraudulent academic world. Most importantly, Dissernet provides a unique view on the deterioration of some institutions' reputations in Russia. In order to illustrate this point, several reference groups may be considered: members of the Russian Academy of Sciences (RAS), principals of Moscow's primary and secondary schools, chancellors of Russian universities, regional governors, and members of the State Duma. Members of each group are targeted if they have been awarded an academic degree in the last 15 years. Dissernet has not detected any falsified dissertation by the RAS members. Of the 141 dissertations defended by principals of Moscow's primary and secondary schools, 23 have satisfied the Dissernet criteria for largely plagiarized texts. This amounts to 16 % — a rate that is more than three times as high as the probability of finding large-scale plagiarism in a random pedagogical dissertation. This implies a silent mechanism at work selecting and supporting those who are prone to falsifications. Another group, chancellors of Russian universities, has shown an even higher fraction of 21 %. Of that, one third of the universities are in Moscow. The proportion of politicians representing regional governors and members of the State Duma is even higher, reaching 41 % for the latter.

DISCUSSION AND CONCLUSIONS

Such large-scale dissertation fraud in Russia is a result of corruption that has paralyzed the whole system of awarding academic degrees (Denisova-Schmidt, 2016): from dissertation councils established by leading universities, to the Higher Attestation Commission, to scientific journals' editorial boards (Osipian, 2012). Dissernet has also traced an apparent involvement of Russian scientific journals in the fake dissertation industry run by universities

(more exactly, certain dissertation councils). Those three pillars (dissertation councils, the Higher Attestation Commission, and journals' editorial boards) are the necessary working parts of the mechanism running the conveyor belts of academic fraud in Russia. Very often, the same people are engaged in these three key bodies in parallel.

It has become obvious that the phenomenon of scientific fraud in Russia is not a marginal one. It is not localized somewhere in the country's hinterland. Today it is playing the role of an institution that is well integrated into the contemporary Russian state. Several recent laws and decrees protect the awardees of falsified academic degrees. The most important one makes it impossible to strip a person off an academic degree if its defense took place before 2011 (RF Government, 2013). The authorities are quite reluctant to revoke fake academic degrees, even if the defense happened after 2011. The reactions of those accused of plagiarism by Dissernet varies from ignoring it, to calling it nonsense, to allegations of being politically motivated, to accusing Dissernet members of incompetence and claiming that only authorised dissertation councils are entitled to assess the dissertations. This point of view is broadly communicated by the state-owned mass media. Still, so far, Dissernet has managed to persuade dissertation councils to revoke about a hundred fake academic degrees.

These studies allow considering the developed fake academic degrees institution as an indicator of the country's scientific and technological weakness. To this end, Dissernet has developed, tested, and offered a technological platform to estimate the scale and reach of fraudulent academic practices. The data obtained could be used as an indicator to rank the country's universities and academic climate.

ACKNOWLEDGMENT

We express our gratitude to hundreds of Dissernet's voluntary participants, mostly professional researchers, that have hugely contributed to our study but decided to remain anonymous. The project is fully non-commercial.

REFERENCES

- The decree of the Government of the Russian Federation from (2013) September 24, #842 Moscow "About the order of award of scientific degrees". Available at <http://rg.ru/2013/10/01/stepen-site-dok.html>
- Denisova-Schmidt, E. (2016). Academic Dishonesty or Corrupt Values: the Case of Russia, in: Torsello, D. "Corruption in Public Administration: an Ethnographic Approach", Edgar Elgar, forthcoming.
- Osipian, A. (2012). Economics of Corruption in Doctoral Education: The Dissertation Market. *Economics of Education Review*, 31, 76-83.
- Rostovtsev, A. (2015). Some Observations on the Subject of Dissertation Fraud in Russia. *HERB: Higher Education in Russia and Beyond*, 3(5), 17-18.
- Weber-Wulff, D. (2014). *False Feathers: A Perspective on Academic Plagiarism*. Heidelberg, Berlin: Springer.

CHAPTER 7

Programme and Project Assessment



An assessment of EU-funded research projects: innovators and their innovative potential¹

Daniel Nepelski^{*}, Annarosa Pesole^{**}, Vincent Van Roy^{**} and Eoghan O'Neill^{**}

^{*} Daniel.Nepelski@ec.europa.eu

European Commission, JRC Institute for Prospective Technological Studies, Calle del Inca Garcilaso 3, 41092 Seville (Spain)

^{**} Annrosa.Pesole@ec.europa.eu; Vincent.Van-Roy@ec.europa.eu; Eoghan.O'Neill@ec.europa.eu

European Commission, JRC Institute for Prospective Technological Studies, Calle del Inca Garcilaso 3, 41092 Seville (Spain)

European Commission, Directorate General for Communications Networks, Content & Technology (DG CONNECT), Brussels (Belgium)

ABSTRACT

The European Commission's Framework Programme constitutes an important share in R&D expenditures in Europe. A number of FP7 projects certainly produce cutting-edge technologies and a significant percentage of these technologies could be commercialized. However, there is a general feeling that not all these technologies and innovations with commercial potential reach the market. The question is why? The Innovation Radar (IR) is a support initiative that focuses on the identification of high-potential innovations in the ICT FP7 projects and the key organization in delivering these innovations to the market. The current paper documents the details of the IR methodology and the results of its first application.

INTRODUCTION

It is widely recognised among economists that R&D and innovation play a key role in driving economic growth and prosperity (Aghion & Howitt, 1992; Griliches, 1979; Romer, 1990). Due to market imperfections, firms are prevented from reaping all the benefits of their R&D efforts. This results in firms investing in R&D below the social optimum (Arrow, 1962; Nelson, 1959). Over the last decades, public support mechanisms have been established by policy makers to encourage firms' investment efforts.

In an attempt to increase Europe's competitiveness in the global knowledge economy, the European Commission launched its seventh Framework Programme (FP7) for Research and Technology Development aiming at financing research grants from 2007 until 2013. This research programme constitutes an important share in R&D expenditures in Europe. Over the whole period, the FP7 has a budget of over €50 billion with €9 billion allocated to ICT (EC, 2007). In comparison, the ICT sector R&D annual expenditures in the EU reached almost €30 billion in 2011 (JRC-IPTS, 2014).

In view of the highly-debated European paradox, it is primordial to appraise the strength of European scientists and entrepreneurs in transforming scientific advances into wealth-generating innovations (Dosi et al., 2006; EC, 1995). To our best knowledge, so far no study has been devoted to the creation of an extensive data infrastructure providing new insights and understanding of the innovations that emerged from the EU-funded research projects along with the innovators associated to these innovations. The current study bridges this gap

¹ Disclaimer: The views expressed are those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

in the scientific literature by proposing the Innovation Radar (IR) which is a DG CONNECT / JRC IPTS support initiative launched in August 2013. The main objective of the IR is to identify the maturity of innovations developed under FP7 projects with an ICT theme and to assess the potential of innovators and innovations. In addition, by collecting information on innovators' needs and bottlenecks, the IR provides guidance on how to offer tailor-made policy support to innovators in EU-funded research projects. This way, the IR represents a policy tool for innovation management and commercialisation.

The IR is constructed from a comprehensive database based on a survey questionnaire that has been administered during standard review processes of ongoing EU-funded FP7 projects under the ICT theme. These reviews are conducted by a panel of independent evaluators, who are recognized specialists in the relevant fields. During their lifecycle, FP7 projects go through three review rounds. Hence, the IR covers and assesses innovations and innovators that belong to projects at different stages of their lifecycle.

The present paper is a pilot study based on 279 FP7 projects in ICT, but is currently scaled to cover all collaborative projects launched under the ICT theme of Horizon 2020. An innovation radar visualisation tool will become available in the near future facilitating policy makers to use the database for policy development. Regular updates of the Innovation Radar after every review round will empower policy makers to quickly respond to the ongoing needs of entrepreneurs conducting EU-funded research projects. In addition, the Innovation Radar can serve as a valuable input for evaluation studies carried out over EU Framework Programmes to assess their impact on European science, research and economic growth (De Prato et al., 2015).

METHODOLOGY

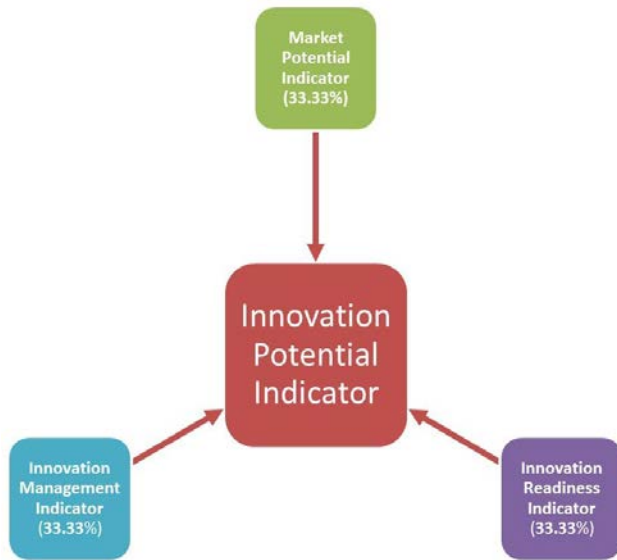
The IR contains two composite indicators aiming at capturing the heterogeneity in innovation activities and innovators across a pilot of 279 FP7 projects in Europe. The first indicator provides a holistic view of the innovation potential of FP7 projects, while the second one is capturing the innovator's capacity in conducting innovation activities. In the remainder of this study these composites are respectively denoted as the "Innovation Potential Indicator" and the "Innovator Capacity Indicator".

a. Innovation potential indicator

In order to define framework for the construction of the "Innovation Potential Indicator", we use three assessment criteria that are commonly referred to in the context of innovation potential assessment exercises: Market Potential, Innovation Readiness and Innovation Management (De Coster & Butler, 2005; Liao & Witsil, 2008). Market potential criterion relates to the demand and supply side of an innovation. Innovation readiness criterion relates to the technical maturity of an evolving innovation. Innovation Management criterion addresses the issue of the project consortium and its commitment to bring an innovation to the market, an element that is often seen as the most important success factor of a technology venture.

In order to observe and measure the above specified criteria, each of them was matched with relevant questions of the Innovation Radar Questionnaire. The IPI is an arithmetic composite indicator which aggregates the values of the three earlier sub-indicators. The construction of the Innovation Potential Indicator is depicted in Figure 1.

Figure 1: Construction of the Innovation Potential Indicator

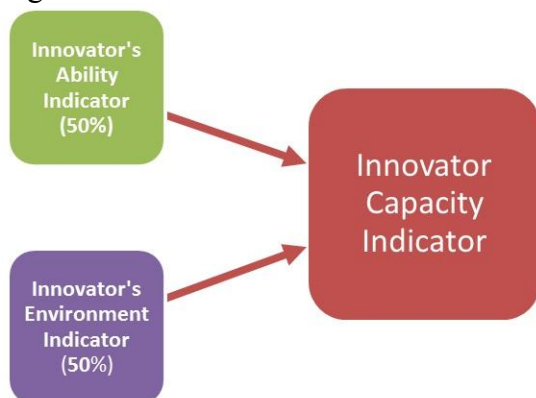


Source: JRC-IPTS

b. Innovation capacity indicator

The innovation capacity indicator aims at unravelling the capacity and capability of innovators in conducting and delivering successful innovations. In line with the absorptive capacity argument that innovators need enough internal capacity to recognize and exploit research from external sources, following criteria were adopted: innovator's ability and innovator's environment (Cohen & Levinthal, 1990; Keller, 1996). Innovator's ability relates to the innovation performance of an individual organization that is seen as the key organization behind an innovation. The innovator's environment criterion aims to capture the overall conditions which an innovator faces. It is mainly related to the overall composition and activity of partner organizations, the performance of the project in terms of innovation and the commitment of relevant partners in exploiting the innovation. Figure 2 visualizes the construction of the Innovation Capacity Indicator (ICI).

Figure 2: Construction of the Innovation Capacity Indicator



Source: JRC-IPTS

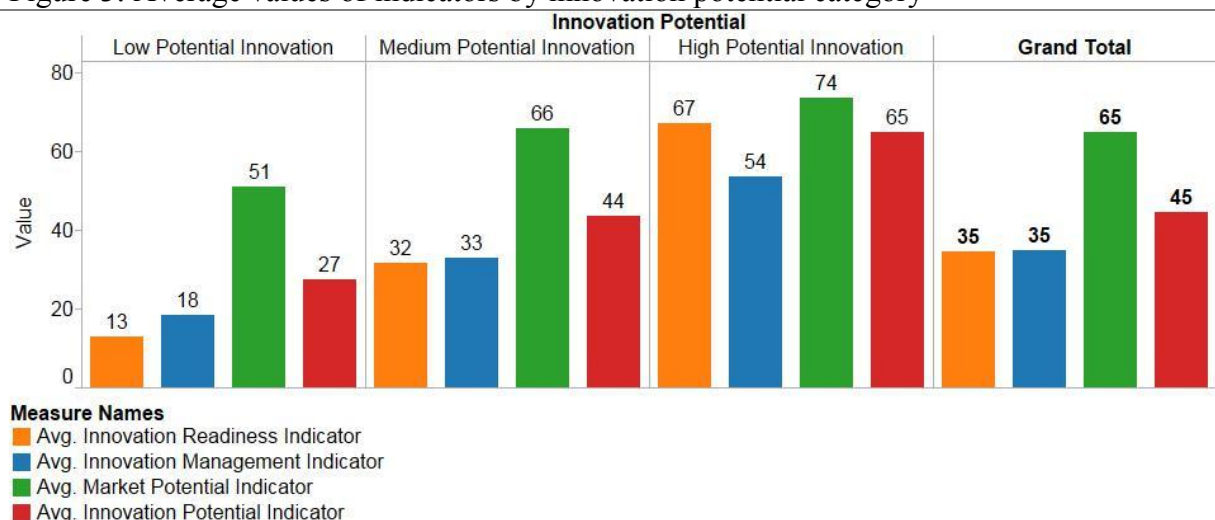
RESULTS

In the following paragraphs, we provide a selection of the results obtained from the Innovation Radar database. In addition to the presented results, the database offers enriching insights on the type of organisations and their innovations, the development stage of innovations, the steps and barriers to innovation commercialisation and the location of innovators (see De Prato et al. (2015)).

a. Overview of innovation performance

Figure 3 reports the average values of the three innovation potential assessment sub-indicators, i.e. Innovation Readiness (IRI), Innovation Management (IMI), Market Potential (MPI) and the composite Innovation Potential (IPI), for all analysed innovations and by innovation potential category. Based on the presented evidence, it can be concluded that, on average, market potential and innovation readiness are among the strongest dimensions of the innovations coming out of the reviewed ICT FP7 projects. In contrast, innovation management represents the weakest dimension of these innovations. Hence, in order to increase the potential of these innovations, steps such as the clarification of innovation ownership, preparing business plan and market study or securing investment must be taken.

Figure 3: Average values of indicators by innovation potential category



Calculations: JRC-IPTS

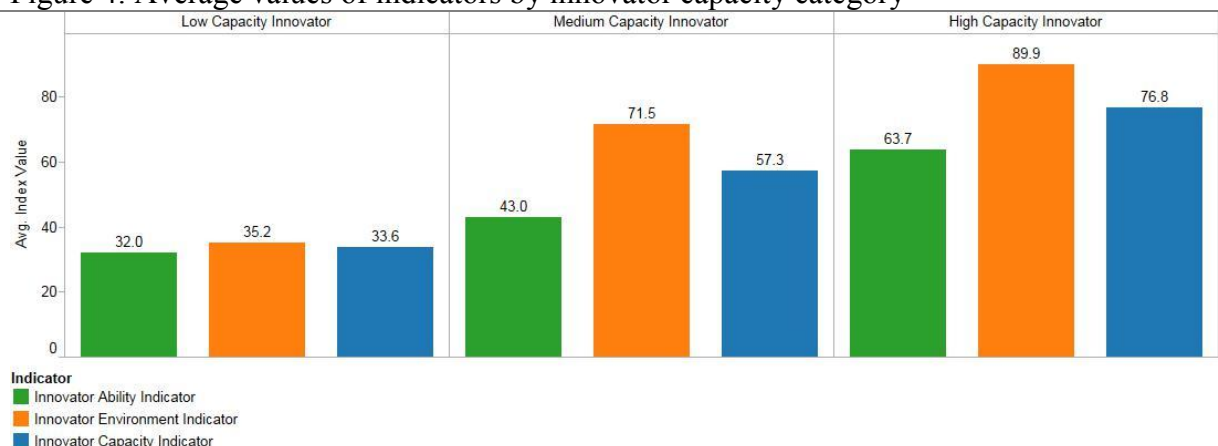
Data: European Commission DG Connect

Note: The figure presents average scores across all four innovation potential assessment indicators, i.e. IRI, IMI, MPI and IPI, across innovation potential categories. Total number of reviewed projects: 279. Total number of innovations: 517. Review period: 20.05.2014 and 19.01.2015.

b. Overview of innovator performance

Figure 4 summarises the average values of the two sub-innovator capacity assessment indicators, i.e. Innovator's Ability (IAI), Innovator's Environment (IEI) and a composite Innovator Capacity Indicator (ICI), for all key organizations in delivering the innovations and by innovator capacity category. Looking at the individual sub-indicators, one can observe that IEI has the highest and the IAI has the lowest average value. The average IEI score is 68.29 and the average IAI score is 44.35 points. However, looking at the standard deviation values of both indicators, one can see that IEI scores are much more volatile than those of IAI. In other words, the quality of the innovation environment is not equal for all innovators. This happens as some projects do not engage end-users in the consortium or because the relevant partners are not sufficiently committed to exploiting the innovation.

Figure 4: Average values of indicators by innovator capacity category



Calculations: JRC-IPTS

Data: European Commission DG Connect

Note: The figure shows the average scores of innovators across three assessment indicators, IAI, IEI and ICI. Innovators are defined as key organization behind the innovation according to the answers to the question 12) of the Innovation Radar Questionnaire. Total number of reviewed projects: 279. Total number of innovators: 544. Review period: 20.05.2014 and 19.01.2015.

Innovation Radar as a business intelligence tool

The Innovation Radar is becoming a tool providing information on the innovation performance of EC-funded projects, innovators' capacity and their needs. This information will be available at various organizational levels. At the operational level, it will help project coordinators to spot potential innovations and the bottlenecks to bring them to the market. This will serve them as a guideline as to what support is needed in order to increase their chances of successful commercialization. At the strategic level, it will help to adapt and design ongoing and forthcoming EC research programs so that they produce desired outcomes in terms of innovative activity and its commercialization. It will also facilitate the provision of policy support to improve the innovation performance and bringing research to the market. This way, the Innovation Radar is expected to become a business intelligence tool. In order to achieve this, it will have to provide real-time data and intelligence that will be made available to policy makers and the EU-funded organizations. The first step towards this is the creation of an Innovation Radar online data visualisation tool (see Figure 5).

CONCLUSIONS

Having completed the pilot application of the Innovation Radar and having analysed the collected information, it can be concluded that, for the first time, policy makers and project coordinators at the European Commission can obtain up-to-date information on the innovative output of these projects. The IR allows them to characterise innovations with respect to their technical readiness, innovation management and market potential. For innovators, it can deliver information on their individual performance and ongoing needs and the environment in which they innovate. The approach pursued under the pilot has now been improved on the basis of lessons learnt in the pilot and is being scaled to cover all collaborative projects launched under the ICT theme of Horizon 2020. A business intelligence dashboard is being developed for EU policy makers to help them make use of these data sets for policy development and to empower a more data-driven approach to managing the Horizon 2020 programme.

Figure 5: Screenshot of the Innovation Radar online visualisation tool (alpha version, under development at time of writing)



Source: European Commission, DG CNECT
Calculations: JRC-IPTS

REFERENCES

- Aghion, P., & Howitt, P. (1992). A Model of Growth Through Creative Destruction. *Econometrica*, 60(2), 323-351.
- Arrow, K. J. (1962). Economic welfare and the allocation of resources for invention. In R. Nelson (Ed.), *The Rate and Direction of Inventive Activity* (pp. 609-626). Princeton, NJ: Princeton Univ. Press
- Cohen, W., & Levinthal, D. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 35(1), 128-152.
- De Coster, R., & Butler, C. (2005). Assessment of proposals for new technology ventures in the UK: characteristics of university spin-off companies. *Technovation*, 25(5), 535-543.
- De Prato, G., Nepelski, D., & Piroli, G. (2015). *Innovation Radar: Identifying Innovations and Innovators with High Potential in ICT FP7, CIP & H2020 Projects*. Seville: JRC-IPTS.
- Dosi, G., Llerena, P., & Labini, M. (2006). The relationships between science, technologies and their industrial exploitation: An illustration through the myths and realities of the so-called 'European Paradox'. *Research Policy*, 35(10), 1450-1464.
- EC. (1995). *Green paper on innovation*. Luxemburg: European Commission.
- EC. (2007). *FP7 in Brief. How to get involved in the EU 7th Framework Programme for Research*. Luxembourg: Office for Official Publications of the European Communities.
- Griliches, Z. (1979). Issues in assessing the contribution of research and development to productivity growth. *Bell journal of economics*, 10(1), 92-116.
- JRC-IPTS. (2014). *The 2014 PREDICT report. An Analysis of ICT R&D in the EU and Beyond*. Seville: European Commission, JRC-IPTS.
- Keller, W. (1996). Absorptive capacity: On the creation and acquisition of technology in development. *Journal of Development Economics*, 49(1), 199-227.
- Liao, P., & Witsil, A. (2008). *A practical guide to opportunity assessment methods*: Research Triangle Park, NC: RTI Press.
- Nelson, R. (1959). The Simple Economics of Basic Scientific Research. *Journal of Political Economy*, 67(3), 297-306.
- Romer, P. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98(5, Part 2), S71-S102.



Evaluating the impact of public space investments with limited time and funds: (methodological) lessons from a Swiss case study¹

Franz Barjak

franz.barjak@fhnw.ch

School of Business, University of Applied Sciences and Arts Northwestern Switzerland FHNW,
Riggenbachstrasse 16, CH-4600 Olten, Switzerland

ABSTRACT

The paper suggests a methodology for evaluating innovation support policies and funding in the space sector. Previous evaluations have suggested methodologies which require considerable time and resources. Our approach combines a data collection at organisational level through standardised interviews and at project level through an online survey which are relatively quick to implement and less costly. We demonstrate that valid results can be obtained with such an approach.

INTRODUCTION

Space research and technological development is one important sector of public research and development (R&D) in many developed countries. The OECD (2014) estimated that in OECD countries in 2013 all national space investments (civil and military budgets) added up to more than 50 billion USD (calculated with purchasing power parities PPP). Almost 80% of this total has been raised by the United States, which is also the country with the highest institutional space budget per inhabitant and year (123 USD PPP). However, other countries have sizable space budgets as well (OECD, 2014).

Public space investments may take on a variety of formats, including internal governmental spending for space technology developments, public procurement for space missions, technological programmes for advancing space technologies, research programmes to generate new fundamental knowledge on space and others. Different government departments (including defence) and agencies fund space activities either directly at national level or through international organisations like the European Space Agency (ESA).

Data availability, costs and timing are frequent limitations and restrictions to implementing the full effects assessment of public space investments. Still, assessing the outcomes of funding programmes and documenting their benefits are important steps in justifying and securing the funding from policy makers and parliaments. Many countries conduct ad hoc impact assessments of the institutional space funding (Simmonds et al., 2012), which then fall short of a more in depth methodological assessments and scientific discussions of the results. Regular evaluations have only been published for Norway and the UK (OECD, 2014). This paper suggests a qualitative methodology for such a regular impact assessment. Its advantages are, compared to other methods:

- Primary data collection is fast and relatively inexpensive for the geographically distributed space community.
- Multi-level analyses of projects, programmes, and organisations are facilitated.

¹ This work was supported by the Swiss Space Office, Swiss State Secretariat for Education, Research and Innovation (SERI).

- Different types of direct and indirect effects can be accommodated, in particular scientific, technological, economic, and social effects.

APPROACHES TO EVALUATIONS OF PUBLIC SPACE INVESTMENTS

The BETA method

The BETA method has focused above all on the measurement of the indirect effects of public R&D programmes. BETA considered effects which are directly related to the objectives of programmes and projects as defined at project start as *direct effects* (Bach & Matt, 2005). For instance, if a project's aim is to build and deliver an instrument for a satellite, the built instrument would be the direct effect. *Indirect effects*, on the other hand, "are derived from application of what has been learnt during the execution of the project, through the activities of the project participants, but which is not directly related to the objectives of the project" (Bach & Matt, 2005, p. 256). The main idea behind indirect effects is to cover and measure the different types of learning that result from funded projects in the funded organisations. Four types of indirect effects have been distinguished (Bach & Matt, 2005, p. 256; Cohendet, 1998; Georghiou, Rigby, & Cameron, 2002): 1) technological effects, 2) commercial effects, 3) effects on organisation and methods, and 4) work-related effects.

The effects of a project are the combined direct and indirect effects which therefore need to be quantified. In extensive face-to-face interviews project participants are questioned about direct effects, the added values of indirect effects, the respective importance of technological, commercial, organization and methods dimensions ("Q1 coefficients") and the estimated influence of public funding, e.g. ESA contracts, on these dimensions ("Q2 coefficients"). All data should be provided as ranges and for the subsequent calculation of effects only the lower boundaries are considered – BETA calls the result therefore a minimum estimate.

The BETA approach combines different types of economic effects and suggests a methodology to measure them. However, it has a strong focus on companies and the business sector. Effects of scientific research are excluded unless their added value can be quantified. New fundamental knowledge that is "only" included in academic publications but not commercialised or implemented in practice is not accounted for. Another problem of the approach is its resource-intensity: interviews have been described to last on average 3 hours involving 2 interviewers and 1-4 interviewees (Georghiou et al., 2002, p. 231). Interviews should be done face-to-face; telephone interviews are not recommended. As the project is the analytical unit all partners to a sampled project have to be interviewed (Georghiou et al., 2002). Moreover, as data collection is on individual projects, the benefit of being able to compare the effects of different projects is obtained for the cost of doing multiple interviews per organization. Letting responding organisations chose the projects on which they report certainly introduces a bias to the more successful projects. However, making a random selection will lead to non-responses whenever organizations are unable or unwilling to report on particular projects. Another problem stems from the sensitivity of the collected data and BETA stressed that the generated data needs to remain confidential and should not be given to the agencies which commissioned the evaluation (Cohendet, 1998). This is a challenging point for evaluation practice.

The Technopolis approach

Commissioned by ESA, Technopolis suggested in 2012 a "Methodology to Evaluate the Direct and Indirect Economic and Social Benefits of Public Investments in Space" (Simmonds et al., 2012). Technopolis made practicable recommendations for ESA to take forward the evaluation of public investments in space. In addition to six economic and six social impact categories the report also considers two categories of environmental impacts. It then goes on

to suggest methodological advancements for each of these 14 impact categories and puts together three different scenarios of different scope for implementing these methodological advancements.

The Technopolis approach includes separate suggestions for data (additional primary data collections and uses of existing databases) and methodological improvements by type of impact and for aggregating impacts. It is a comprehensive approach for assessing public space investments which still needs considerable research and development activities in order to become feasible (which the authors estimate in the range of 3-5 mEUR, Simmonds et al., 2012).

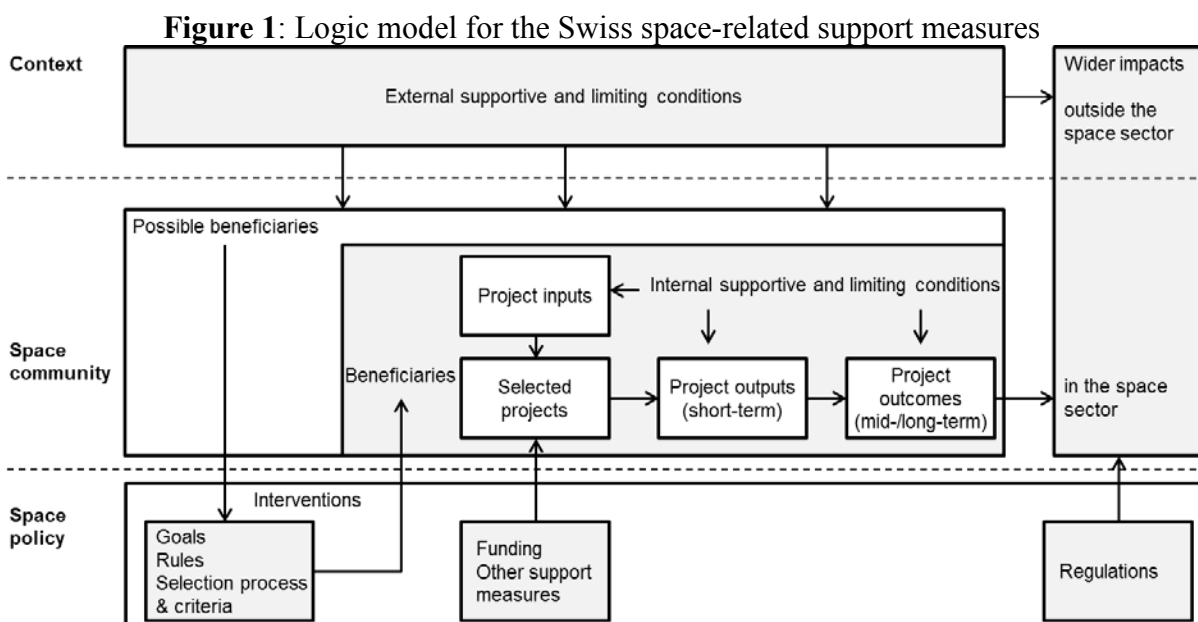
SCBA-plus

A "top-down" approach for measuring the impacts of space funding was suggested in Hof, Koopmans, Lieshout, and Wokke (2012). They advocated a Social Cost Benefit Analysis (SBA) combined with a Multi Criteria Analysis (MCA) and baptised this "SCBA-plus". All effects that can be measured and monetised should be valued according to the SCBA method which will generate an overview of effects in their own terms and in money terms. All other effects should be measured by using MCA and summarised in their own terms and in the form of scores from a uniform score card. SCBA-plus also requires detailed data on space investments, the space sector, and sales of spin-off technologies plus an assessment of the non-quantifiable (social, strategic & environmental) effects.

Our approach differed from these three approaches by requiring less time and resources, of course with the negative effect of gaining fewer insights into the consequences of space funding.

APPROACH CHOSEN FOR THE IMPACT ASSESSMENT OF PUBLIC SPACE INVESTMENTS IN SWITZERLAND

As others before us we developed a logic model that related public space investments to the effects, taking into account other possible influences on the effects as well (see figure 1). It distinguishes between the space community (middle part), the context (upper part), and space policy (lower part).



The space community according to this definition consists of the possible beneficiaries of space-related funding, which are business enterprises or companies and research institutes engaged in space-related activities (OECD, 2012). Governmental actors were excluded as they are usually not entitled to receiving space funding from public sources. Recipients for space funding are selected from these possible beneficiaries, through applications and contracts and according to defined processes, rules, and criteria. These beneficiaries receive the public funding for a specified activity and purpose, which are usually defined in the underlying programme and in the project itself. However, whether project goals can be achieved depends not only on the funding, but also on other internal inputs and conditions as well as external inputs and conditions. These internal and external inputs need to be captured as well in order to identify the funding effect.

Project outputs refer to the products generated by space-related projects, e.g. the knowledge, technologies, processes, instruments, infrastructure, services etc. directly invented, developed or built by project participants with the funding. Outputs will often reflect projects' goals or objectives, but not only in theory might unintended outputs, coincidental insights, or serendipitous discoveries appear as well. Outputs are heterogeneous and usually lack a price tag or common measure that permits for easy aggregation. Evaluations therefore tend to follow one of two approaches for measuring outputs:

- Equating output values with input values, i.e. measuring total project budgets including public funding and funding from other sources (Danish Agency for Science Technology and Innovation, 2008); this approach, however, essentially fails to capture the value that is generated within the projects.
- The other approach is qualitative: it presents and (eventually) counts the different outputs generated or advanced through the funding, e. g. scientific publications, new technologies, patents and licences etc. (Academy of Finland, 2004; Amesse, Cohendet, Poirier, & Chouinard, 2002; Centre for Strategy & Evaluation Services (CSES), 2011). This approach lacks a uniform measure for the different outputs, complicates comparisons across programmes and projects, and does not provide monetary values. Still, it is currently the best approach to capture the diversity and richness of outputs.

As project outcomes we defined all those effects that result among the participants of funded space-related projects due to the project realisation very much in line with the BETA definition (see above). These outcomes can translate into an increase in sales, cost reductions and/or the existence of a critical mass of space-related specialists.

Wider impacts of space-related funding cover in our understanding those which go beyond the participants in funded projects and reach industry, academic organisations, or society as a whole. The measurement of wider impacts is for many reasons challenging (Clark et al., 2014; Cohendet, 1998; OECD, 2012): identifying all impacts and recipients, attributing causality to the space-related funding, putting a price tag or even quantifying impacts require insights and data which are not readily available. In addition, effects will not only be direct, but also indirect. General impact evaluations therefore have at most listed the social and environmental effects which could be identified by program participants and beneficiaries of the funding. This can only be considered as anecdotal evidence of limited value. Anything more sophisticated, however, is beyond impact assessments of space funding and only possible in the framework of dedicated studies (Clark et al., 2014).

THE SETTING OF THE SWISS CASE STUDY

Swiss space policy and funding

The Swiss Federal Dispatch on Education, Research & Innovation foresees a total of mCHF 528.2 (approximately 480 mEUR as of October 1st, 2015) of space-related funding for the time period 2013-2016 (Schweizerischer Bundesrat, 2012). The main goals of this funding and Swiss endeavours in the space sector are according to the Swiss government (Swiss Confederation, Federal Department of Economic Affairs, Education, and Research, 2013): 1) To secure access to space; 2) develop and use space applications to improve the quality of life for citizens; 3) promote existing and new focal areas of Swiss excellence and innovation and allow entry to the critical technological path in the development of space infrastructure; 4) and contribute to Switzerland's reputation as a scientifically powerful and technologically innovative country.

Switzerland pursues these goals through funding channelled through mainly two instruments: 1) Contributions to the European Space Agency (ESA) allowing participation to a large number of its programmes, and 2) funding of the complementary national activities (Activités Nationales Complémentaires ANC) implemented through the Swiss Space Office (SSO). The latter are implemented in order to improve the Swiss scientific and technological position mainly in European space programmes and space industry. The funds are used for funding of selected infrastructures and key projects in specific domains, in particular favouring technology transfer from academia to industry.

Setting of the evaluation

This article draws on an impact analysis conducted for the Swiss Space Office. The study responded to a call for a qualitative assessment of the effectiveness of Switzerland's participation to ESA programmes and the ANC investments. The call asked for an analysis of the relevance, effectiveness, impact, sustainability, and coherence of the funding. A time period of 4.5 months and a corresponding budget for 2 FTE were foreseen to conduct the impact analysis. During these 4.5 months desk research, the collection and analysis of data and the writing of the report had to be terminated. The contractor offered support with approaching the funded organisations and developing the data collection instruments.

The following section describes the methodology as suggested by the evaluators to conduct the impact analysis.

Evaluation design

Approach

The evaluation approach was subject to several restrictions:

- Due to timing and funding of the impact analysis, the lack of well-established measures, the impossibility to realise a dedicated data collection and reach out to beneficiaries and users in the corporate sector, administration, the health sector, or society in general, an assessment of the wider impacts of space funding were beyond the scope of the study. It had to be limited to beneficiaries and participants in space projects.
- The Swiss space sector is overall rather small with less than 200 organisations. As nearly all organisations also participated in funded projects it was impossible to assemble a control group of unfunded entities within the same sector.
- Space funding has been provided for a number of years and a before-after comparison was neither possible.

We opted for a survey design that asked questions to all Swiss space organisations (using different survey modes) and on a sample of the realised space projects.

Methods of data collection

In order to answer the questions of the call data on three units of observation was needed: a) Policy interventions, i.e. the funding programmes for space-related activities, b) funded space projects, and c) the organisations conducting these space projects. While some data might have been collected with comparable quality at any of these three levels, other data was level-specific; e.g. funding regulations and procedures are programme-specific, whereas the goals and outputs are project-specific. We covered all three levels in the data collection and combined projects and programmes. The data collection relied on three methods:

- 1) *Semi-structured face-to-face interviews* with organisations have been used in previous evaluations to collect data on the impact of space-related funding (Bach, Cohendet, & Schenk, 2002; Cohendet, 1998; Danish Agency for Science Technology and Innovation, 2008; Hertzfeld, 2002; Prognos AG, 2008). Face-to-face interviews were held in this evaluation with twenty key players of Swiss space industry and space research. They were recorded and transcribed verbatim.
- 2) *Structured telephone interviews* targeted Swiss companies and institutes with minor and potential future involvement in the space sector. Interviews were recorded, but only summary transcripts were prepared.
- 3) An *online survey* collected data on selected projects of the interviewed organisations (project inputs, goals, outputs, and outcomes) and on the corresponding programmes.

Combining the three data collection methods had a number of advantages which we considered as crucial for conducting the analysis successfully:

- Organisations which had experiences with several and bigger projects had been funded for many years and presumably knew ESA and ANC funding opportunities and conditions well. They could provide rich information on the funding and its context which needed face-to-face interviews with some flexibility.
- The telephone interviews were sufficient for organisations for which the space business was only a side business. The survey mode also permitted to approach a geographically distributed set of respondents in a short time period.
- Project-related questions partially needed input from project participants like principal investigators which required that the contact persons in every organisation could delegate the answers. In addition, questions on numbers, e.g. total project budget, might have required asynchronous answering (e.g. in order to check first project documents before answering).

Samples and responses

The samples consisted of 1) Swiss organisations in the space sector and 2) their space projects.

1) Organisations. Through the ESA Yellow Pages, internal information from the SSO and internet searches we identified 153 separate organisations in the Swiss space sector of which one third were academic and two thirds corporate. These organisations were then invited to participate in the survey. During the survey 11 organisations had to be dropped from the sample, because they were unreachable, economically inactive or they had dropped the space business since then and were unable to report on it.

The field phase, originally planned to last 5 weeks, had to be extended to 13 weeks for different reasons. In total 79 interviews were conducted in this period, of which 40% with academic institutes and 60% with companies. The overall response rate was 55.6%, nearly 70% among academic institutes and approximately 50% among companies (see Table 1). 20 interviews were conducted face-to-face (average interview duration of 70 minutes) and 59 by phone (average duration 34 minutes).

Table 1: Realised sample by sector and funding status

	Academic institutes		Companies		Total	
	N	in %	N	in %	N	in %
Cleaned gross sample	47	100.0%	95	100.00%	142	100.00%
Respondents	32	68.1%	47	49.5%	79	55.6%
with space-related public funding (ESA/ANC) 2010-14	26	55.3%	42	44.2%	68	47.9%
without space-related public funding 2010-14	6	12.8%	5	5.3%	11	7.7%
Rejections	7	14.9%	5	5.3%	12	8.5%
Non-responses	8	17.0%	43	45.3%	51	35.9%

Projects. The organisations in the sample had more than 1'000 different ESA contracts and 31 ANC projects. We included all 31 ANC projects and made a systematic selection of 397 ESA-funded contracts for the project-related survey. The field phase of this survey also lasted for 13 weeks. Overall we received 154 filled in questionnaires answered by 18 academic institutions and 31 companies. 98 questionnaires were filled in by companies, 56 by academic institutions. The overall response rate was 36% (see Table 2).

Table 2: ESA and national space-related contracts to Swiss organisations

	All contracts per organisations		Gross sample		Responses		
	N	in %	N	in %	N	in %	Response rate
ESA-funded*	1526	98.0%	397	92.8%	135	87.7%	34.0%
ANC projects	31	2.0%	31	7.2%	19	12.3%	61.3%
Total	1557	100.0%	428	100.0%	154	100.0%	36.0%%

* "Contract" refers to the participation of a Swiss organisation in an ESA project.

LESSONS LEARNED

The results of the evaluation are not reported in this paper (see Barjak, Bill, & Samuel, 2015) which concludes with a few methodological lessons learned.

1) The evaluation was started with a very ambitious time frame. The selected methods of data collection and analysis permitted a successful implementation of the evaluation within little more than the available time frame. The evaluation was successful with producing input measures for space organisations, as well as output and outcome indicators.

2) The approach received rather positive feedback from the organisations contributing the data. However, the aim of generating information on a systematic sample of projects was not fully successful. Several respondents were unable to recognise projects conducted (presumably) by their organisations and project participants could not be identified (e.g. in the case of universities with one organisational ID for several institutes). Hence, the original aim of putting together an unbiased sample of funded projects for the data collection could not be realised and several respondents had to be asked to select projects on which they wanted to respond. This probably introduced a bias towards more productive and more successful

projects. As we do not have any information on 90% of all funded projects we interpreted the project-related results only with great care.

REFERENCES

- Academy of Finland. (2004). *Research Programme for Space Research ANTARES 2001-2004. Evaluation Report*. Helsinki: Academy of Finland.
- Amesse, F., Cohendet, P., Poirier, A., & Chouinard, J.-M. (2002). Economic Effects and Spin-offs in a Small Space Economy: The Case of Canada. *The Journal of Technology Transfer*, 27(4), 339-348.
- Bach, L., Cohendet, P., & Schenk, E. (2002). Technological Transfers from the European Space Programs: A Dynamic View and Comparison with Other R&D Projects. *The Journal of Technology Transfer*, 27(4), 321-338.
- Bach, L., & Matt, M. (2005). Twenty years of evaluation with the BETA method: some insights on current collaborative ST&I policy issues. In P. Llerena & M. Matt (Eds.), *Innovation policy in a knowledge-based economy* (pp. 251-281). Berlin: Springer.
- Barjak, F., Bill, M., & Samuel, O. (2015). *Evaluation of the existing Swiss institutional R&D funding instruments for the implementation of the space-related measures*. Retrieved from <http://www.sbf.admin.ch/themen/01371/01510/index.html>.
- Centre for Strategy & Evaluation Services (CSES). (2011). *Interim Evaluation of FP7 Space. Final Report*. Sevenoaks, United Kingdom: CSES
- Clark, J., Koopmans, C., Hof, B., Knee, P., Lieshout, R., Simmonds, P., & Wokke, F. (2014). Assessing the full effects of public investment in space. *Space Policy*, 30(3, Part A), 121-134.
- Cohendet, P. (1998). Evaluating the industrial indirect effects of technology programmes: the case of the European Space Agency programmes *Policy Evaluation in Innovation and Technology: Towards Best Practices* (pp. 189-223). Paris: OECD.
- Danish Agency for Science Technology and Innovation. (2008). *Evaluation of Danish Industrial Activities in the European Space Agency (ESA). Assessment of the economic impacts of the Danish ESA membership*.
- Georghiou, L., Rigby, J., & Cameron, H. (Eds.). (2002). *Assessing the Socio-economic Impacts of the Framework Programme*. Manchester: University of Manchester Policy Research in Engineering Science and Technology (PREST).
- Hertzfeld, H. R. (2002). Measuring the Economic Returns from Successful NASA Life Sciences Technology Transfers. *The Journal of Technology Transfer*, 27(4), 311-320.
- Hof, B., Koopmans, C., Lieshout, R., & Wokke, F. (2012). *Design of a Methodology to Evaluate the Direct and Indirect Economic and Social Benefits of Public Investments in Space*.
- OECD. (2012). *OECD Handbook on Measuring the Space Economy*: OECD Publishing.
- OECD. (2014). *The Space Economy at a Glance 2014*: OECD Publishing.
- Prognos AG. (2008). *Evaluierung der österreichischen Beteiligung an ESA Wahlprogrammen sowie des österreichischen Weltraumprogrammes des BMVIT*.

Schweizerischer Bundesrat. (2012). *Botschaft über die Förderung von Bildung, Forschung und Innovation in den Jahren 2013–2016*. Schweizerischer Bundesrat Retrieved from <http://www.admin.ch/opc/de/federal-gazette/2012/3099.pdf>.

Simmonds, P., Clark, J., Knee, P., Stermsek, M., Horvath, A., & Javorka, Z. (2012). *Design of a Methodology to Evaluate the Direct and Indirect Economic and Social Benefits of Public Investments in Space*. Retrieved from: <http://www.technopolis-group.com/wp-content/uploads/2014/06/ESA-Evaluation-Methods-Study-Full-Report.pdf>

Swiss Confederation, Federal Department of Economic Affairs, Education, and Research. (2013). *Swiss Space Implementation Plan within Education, Research and Innovation for 2014-2023*. Bern.



Researchers and Institutions in the Periphery: Challenges in Measuring Research Capacity for Geographically Specific Programs in the U.S.¹

Julia Melkers

jmelkers@gatech.edu

School of Public Policy, Georgia Institute of Technology, Atlanta, GA 30332 (USA)

ABSTRACT

The U.S. research funding system supports two geographically targeted government programs designed to enhance the research capacity of less competitive states in seeking federal research funds. Program eligibility is based on a relatively narrow measures of overall jurisdiction research funding track record. Yet, in order to adequately measure and address improved capacity development in these areas, a more nuanced understanding of the research capacity in these settings is needed. While the metric of prior research funding metric may be useful in assigning jurisdictions program eligibility, they do not account for variations within or across these areas, nor for factors that may function differently in lower capacity regions. This paper addresses the particular institutional and social/human capital aspects of measuring and identifying research capacity in these regions. Institutional data are used to categorize institutions within and across eligible jurisdictions. Survey data are used to address capacity issues and factors relevant to capacity development at the researcher level.

INTRODUCTION

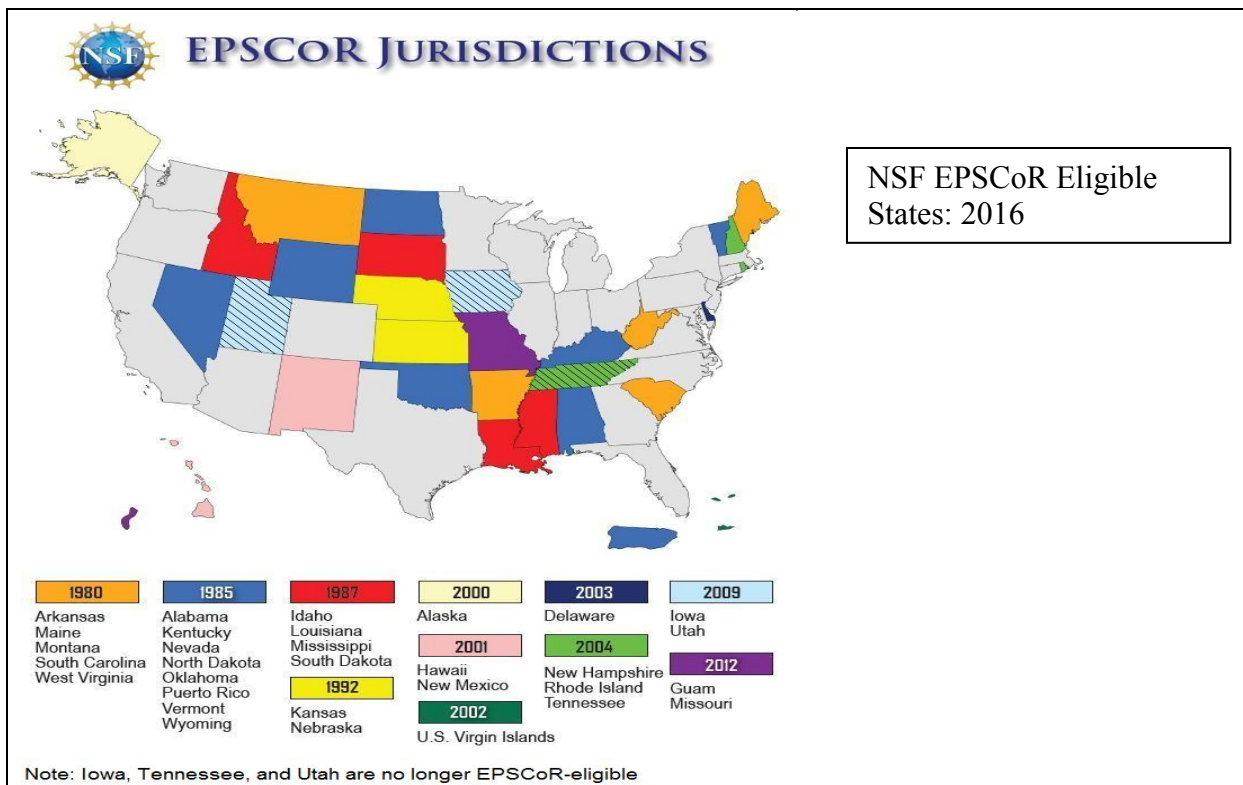
Regional and institutional variation in S&T capacity presents challenges for the structure and management of science funding programs. Overall national research competitiveness will benefit from enhancing the competitiveness of lower capacity regions and institutions in order to create a stronger system overall. In the U.S., certain policy mechanisms have used relatively narrow indicators of research capacity to identify and support researchers and institutions that have been less successful in research grant attainment. More specifically, two geographically targeted government programs designed to enhance the research capacity of less competitive states in seeking federal research funds have been in place for some time. Yet, in order to adequately measure and address improved capacity development in these areas, a more nuanced understanding of the research capacity in these settings is needed. While the metric of prior research funding metric may be useful in assigning jurisdictions program eligibility, they do not account for variations within or across these areas, nor for factors that may function differently in lower capacity regions. This paper addresses the particular institutional and social/human capital aspects of measuring and identifying research capacity in these regions. In the language of the 2016 STI Conference, these regions are peripheral in terms of not only research funding success, but also in factors that matter for this success.

¹ This work was supported by the National Science Foundation (PI Julia Melkers, CoPIs: Eric Welch and Monica Gaughan) (NSF Grant # DRL-0910191).

BACKGROUND

Research funding programs to enhance capacity are abundant. Attention to building increasing interest and retention in STEM disciplines, building a strong STEM workforce, enhancing capacity of early research and underserved populations are evident in many funding portfolios in the U.S. funding agencies. However, in the U.S., there are only two federal *jurisdiction specific national research funding programs* that target the enhancement of research and STEM workforce capacity development in areas that have been identified as less competitive.

The Experimental Program to Stimulate Competitive Research (EPSCoR) which awards most of its funds through the National Science Foundation, and the Institutional Development Award (IDeA) program which awards funds through the National Institutes of Health. Both the EPSCoR and IDeA programs use a jurisdictionally targeted approach to developing research capacity through a focus on “scientific and technical human capital”, which includes both human capital endowments such as formal education and training, and social relations and network ties (Bozeman and Corley, 2004). Using EPSCoR and IDeA as a foundation, this paper addresses the evidence of research capacity development in these less advantaged regions.



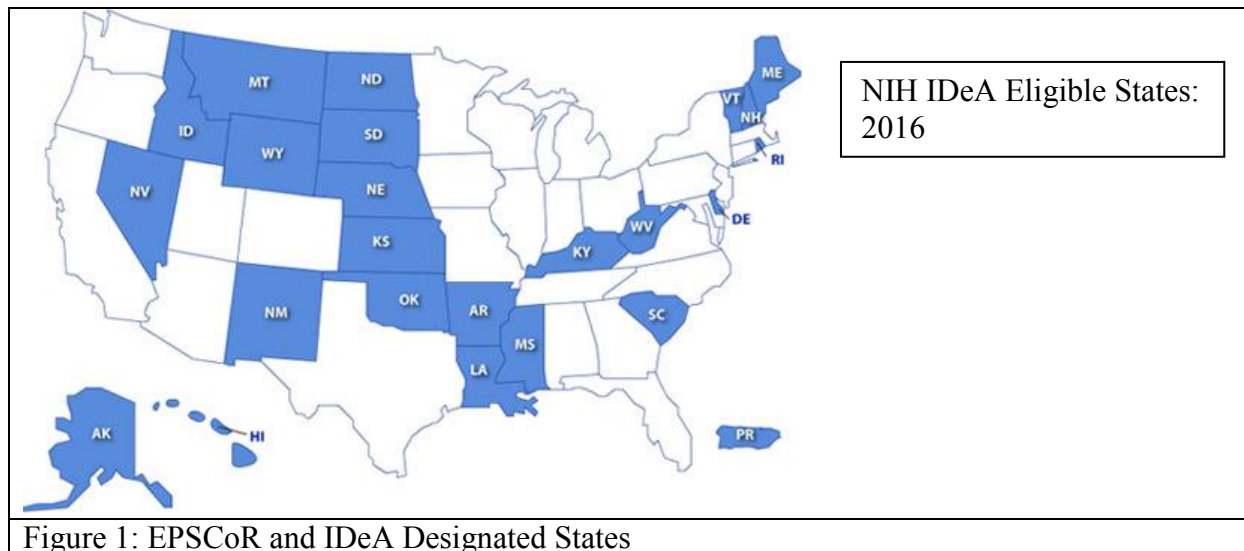


Figure 1: EPSCoR and IDeA Designated States

The EPSCoR program, initiated in 1980, currently funds projects up to \$20 million over five years. It was originally considered to be experimental because it was targeting non traditional institutions and jurisdictions with major research awards in order to catalyse and spur research success (Lambright, 2000.) The NIH IDeA program, begin in 1993, is focused on biomedical research areas and provides funds for five year projects at \$2.5 million per year in direct costs.

For both, researchers and institutions are eligible for these awards if their jurisdiction's overall NSF or NIH funding success (respectively) falls under a particular threshold. Improvements in state's S&T capacity are demonstrated through increases in this share. Both programs seek to build jurisdictional capacity through the engagement of not only the core research institutions(s) in the state, but also through support to smaller regional schools not typically active in federally active research activities. Notably, both programs specifically address state wide research capacity as an expectation of their awardees. This aspect distinguishes these programs as not only addressing lower capacity jurisdictions, but also institutions within those states that have less engagement in the national research enterprise.

RESEARCH QUESTIONS

Specifically, the following questions are addressed: 1) *what are the key indicators of institutional variation, by institutional type, across EPSCoR and IDeA jurisdictions, and compared to noneligible jurisdictions? To what degree do these variations reflect variation in research capacity? How well do they align with program eligibility?* 2) *In what specific ways are scientists in EPSCoR/IDeA states disadvantaged in ways in which EPSCoR/IDeA may better target research capacity building investments?* To answer the first question, we examine the differences between institutions within and across EPSCoR and non EPSCoR states in institutional resources and characteristics. To address the second question, we use researcher reported institutionally based scientific and technical human capital. In this way we are also able to distinguish researchers who are able to develop capacity in these less advantaged regions. The research presented here builds on a prior examination of EPSCOR states (Melkers and Wu,2009) but expands this to include those institutions not traditionally engaged in research, and researchers at all institutions who have been supported through the programs of interest.

Data

The data for this paper comes from both primary and secondary sources. First, in order to develop a better understanding of institutional resources and differences, a comparison of across similar institutions in EPSCoR/IDeA states and those not eligible for either program will be done using U.S. Integrated Postsecondary Education Data System (IPEDS), and the NSF WEB CASPAR data. These sources provide detailed aspect in a broad range of institutional variables and have been generally untapped in an examination of EPSCoR/IDeA. Together these sources will enable the development of a robust understanding of institutional variations within and across EPSCoR/IDeA and other states.

Second, to address issues at the researcher level, this paper will use data from the National Science Foundation funded NETWISE II project, a nationally representative survey of tenured and tenure track faculty in the fields of mathematics, civil engineering, biology and biochemistry. These disciplines were selected to reflect low, transitioning, and higher levels of female representation among faculty, respectively. Further, the sample was constructed to focus on both the research centric institutions, as well as those traditionally less involved in the research enterprise (and often excluded from studies of research activities). The institutional types included here account for nearly 28% of all institutions of higher learning in the U.S., and nearly 75% of all 4 year institutions.

The survey included items on individual background, career path, productivity, teaching activities, professional activities, and professionally relevant social network data. It had a total unweighted response rate of 42%, with 4,195 and includes faculty from 477 institutions. About one fourth of respondents are from EPSCoR/IDeA states, and includes individuals who have been directly supported by these programs.

ANALYSIS

Analysis will include descriptive statistics and comparison of key institutional variables drawn from the IPEDS and WEB KASPARS data across institutional types, and across U.S. states by program eligibility as well as other factors. Survey data analysis will include descriptive statistics and explanatory regression models to address research capacity at the researcher level, as well as factors that differentiate capacity development particularly in EPSCoR/IDeA states.

CITED REFERENCES

- Bozeman, B., J.S. Dietz and M. Gaughan (2001). Scientific and technical human capital: an alternative model for research evaluation. *International Journal of Technology Management* 22(7-8), 716-740.
- Lambright, W. H. (2000). Building state science: the EPSCoR experience. In J. S. Hauger & C. McEnaney (Eds.), *Strategies for competitiveness in academic research*. Washington, DC: American Association for the Advancement of Science.
- Melkers, J., & Wu, Y. (2009). Evaluating the improved research capacity of EPSCoR states: R&D funding and collaborative networks in the NSF EPSCoR program. *Review of Policy Research*, 26(6), 761-782.



Assessing marine biotechnology research centres in peripheral regions: developing global and local STI indicators¹

Douglas K. R. Robinson^{*}, Antoine Schoen^{*}, Patricia Laurens^{*}, Süzel Horellou^{**}, Pierre Colas^{**} and Philippe Larédo^{*}

^{*} douglas.robinson@esiee.fr ; antoine.schoen@esiee.fr ; patricia.laurens@esiee.fr ; philippe.laredo@enpc.fr
LISIS - Laboratoire Interdisciplinaire Sciences Innovations Sociétés (CNRS/ESIEE/INRA/UPEM), Université Paris-Est, Cite Descartes, 5 Bd Descartes 77454 Marne-la-Vallée Cedex 2, France

^{**} suzel.horellou@sb-roscoff.fr ; pierre.colas@sb-roscoff.fr
Station Biologique de Roscoff, Place Georges Teissier, 29680 Roscoff, France

ABSTRACT

Our study tackles the challenge of developing STI indicators for assessing marine biotechnology (Blue Bio) research institutes that are geographically located in peripheral regions, far from major metropolitan areas. The promise of Blue Bio couples (a) the promise of new sources of knowledge and innovation with (b) the promise to stimulate jobs and growth² in regions which struggle to prosper due to a number of factors (such as economic migration from peripheries to large cities, decline of traditional coastal economic activity etc.). In this paper we outline the context of Marine Biotechnology assessment, the framework that is being used, and the first results of its application.

INTRODUCTION

Marine biotechnology, otherwise known as "blue bio", promises advances in biotechnology driven advances in medicine, cosmetics, nutraceuticals, food production, advanced aquaculture, bioremediation and bioenergy. It has been earmarked as a key technology driver for growth, particularly for those countries which leverage large coastlines for economic development (Ritchie et al. 2013).

The majority of countries developing their marine biotechnology capacity have a historical tradition of a marine or maritime economy through marine bioresource exploitation and management (fishing, seaweed harvesting and shipping) or marine research (biodiversity, ecological studies etc.) and it is around these existing hubs that marine biotechnology hubs have been springing up.

Whilst there is some history in marine biology, and now marine biotechnology, research and innovation, many of the research centres are based on the coastal peripheries of their countries – meaning close proximity to the sea, but greater distance from metropolitan areas that are the usual locations of knowledge intensive research and innovation activities. Thus, there is a hope that marine biotechnology will not only provide new knowledge for a variety of industries, but will stimulate jobs and growth in peripheral regions.

It is in this “setting” that our research is located. How does an excellent research centre located in the periphery assess its impact? What are the particular issues for assessing Marine

¹ This work was supported by the EC Horizon 2020 project European Marine Biological Research Infrastructure Cluster (EMBRIC)

² And more broadly, what is described in the theme of the STI 2016 conference as socio-economic transitions in geographical regions.

Biotechnology (Blue Bio) research centres which are often located on the coasts and often far from the main metropolitan centres?

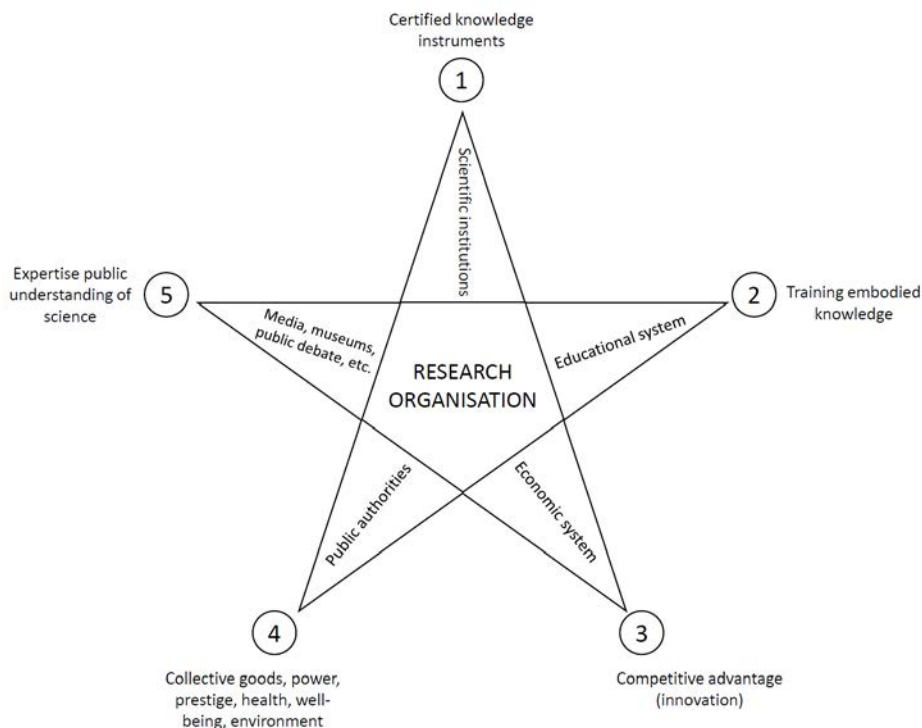
To answer these questions, assessment tools are necessary to enable capacity building, and to promote jobs and growth. More specifically, assessment tools must be developed and applied to (a) capture the existing impact profiles of the different marine biotechnology research centres in terms of scientific excellence and socio-economic impacts and to (b) create comparative criteria for developing best practices and to evaluate, compare and monitor change.

As a pilot study, we choose Roscoff SBR to build and test assessment tools. Roscoff is a research centre *physically* located in periphery region (Brittany), institutionally located in a metropolitan hub (Paris).³ In this *Research in Progress* paper, we will describe our pilot study which we use to develop the assessment approach that will be further applied to approximately twenty other marine biotechnology research centres in Europe as part of a European Horizon2020 project EMBRIC.

THE BROAD FRAMEWORK

For this case, we build a multi-criteria assessment tool to assess the regional embedding of the research centre and the global connections (in the global world of knowledge, but also in national and international networks and links). We mobilise the “research compass” Mustar and Laredo (2000) to position our own descriptors, markers and indicators and assess the research centre in a dual movement of local and global assessment.

Figure 1: An adapted research compass card (Laredo and Mustar 2000)



³ SBR is both a CNRS research centre (French national research organisation for fundamental research) and part of the UPMC (Université Pierre and Marie Curie), with its main campus in Paris.

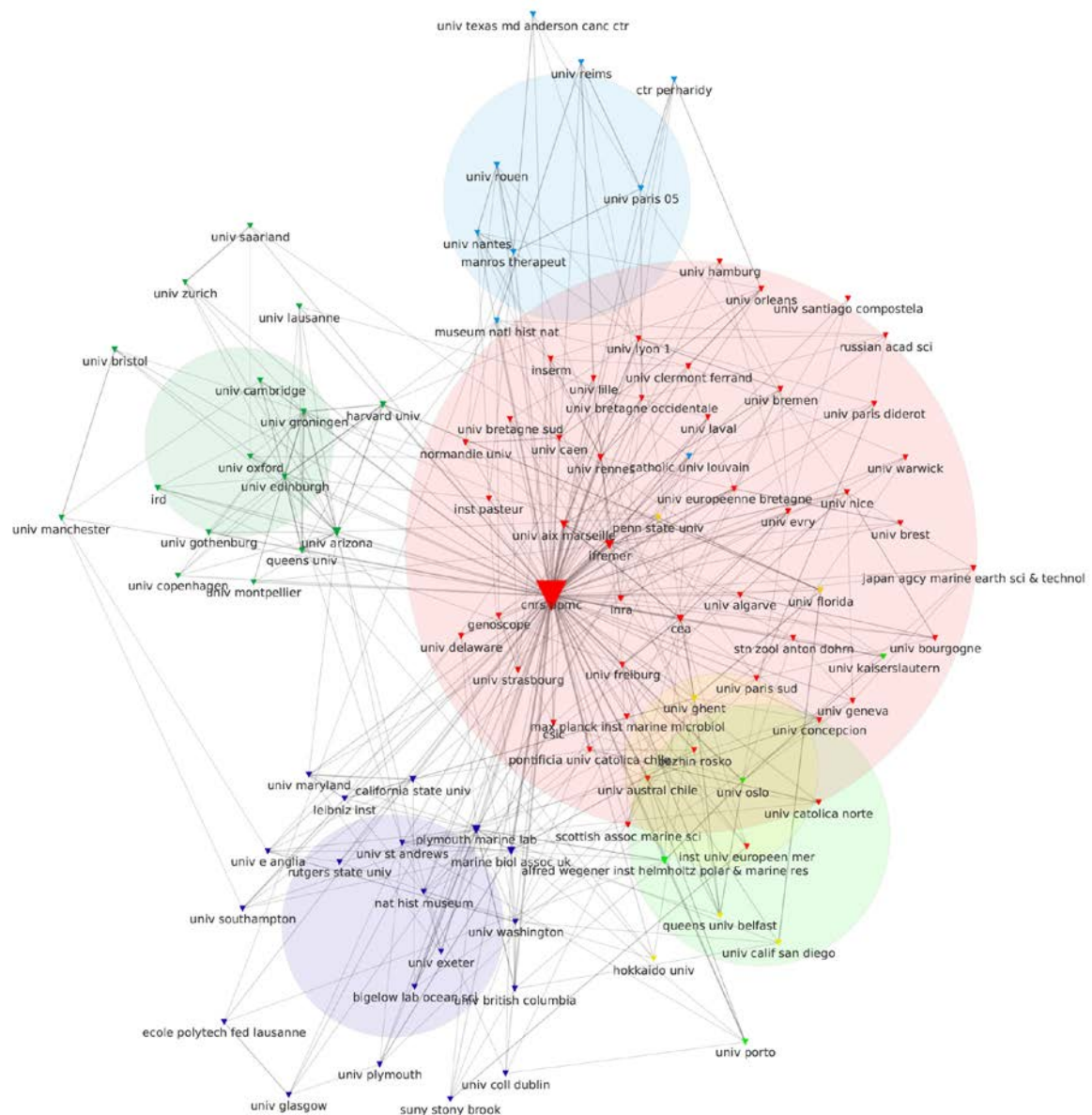
Table 1. Brief description of each point of the research compass

Compass point	Description
(1) Research	The production of scientific knowledge in various forms.
(2) Training	The education of undergraduate and masters students through academic and professional training programmes
(3) Economic activity	Contracts with other actors, such as consultancies, collaborations, joint patenting etc.
(4) Links with public actors	The connection between the research centre and regional public councils and other public organisations. For example, the presence of experts from the research centre on committees, the links between regional development (for example the building of a science district) through interactions with public actors.
(5) Links with society	Links to regional actors and civil society in three broad ways (a) through public engagement and communication, (b) stakeholder interactions to bring to light the norms and values inherent to the research centres activities and how they align with societal actors (Marris et al. 2008)

INITIAL FINDINGS

During this research activity we will develop descriptors, markers and indicators we mobilise for all the points of the compass, however we can already say something with our initial findings. To explore the scientific profile of the research centre, point (1) of the compass, we gathered the publication list of the different units in SBR for the period 2010-2014 inclusive. From this list we extracted from the Web of Science (WoS) the meta-data for all those peer-reviewed articles that could be found in the WoS. From this data we could visualise the cognitive landscape of the corpus, as well as the institutional connections through co-author linkages (see Figure 2 below). From this data we can see that the largest cluster of institutions (see central circle, figure 2) contains both regional and international institutions.

Figure 2: Map of co-author institutions from the 2010 – 2014 SBR corpus of articles extracted from the Wbe of Science (powered by CorText).



To explore the economic/innovation activity, point (3) in the compass, we are currently gathering and organising data on all the contracts between SBR (CNRS-UPMC) and other actors (public agencies, firms, charities and foundations). At the time of writing we have over a hundred contracts in the database where the majority (approximately 90% of private sector contracts are from the region of Brittany.

Considering the public policy linkages too, point (4) of the compass, Public policy links: Its attractiveness for the region has been (and continues to be) recognized by the national Future Investments Programme (PIA - *Plan des investissements d'avenir*). SBR has three PIA projects: the "National Resource Centre for Marine Biology" (EMBR-C-France), which is both a National Infrastructure in Biology-Health (Future Investments Programme) and an

Infrastructure for Research in Agricultural Sciences , ecological and environmental (Roadmap IR DGRI); and "IDEALG" programs and "Océanomics", both in the "Biotechnology and Bioresources" of PIA. Moreover, the regional policy actors are also involved in direct investments into BlueValley (the creation of a Blue Bio science and innovation district in the Finistere region of Brittany) and BlueTrain (large scale structuring and development of the regions academic and professional training in Blue Bio).

GOING FORWARD

Our research seeks to distinguish between “loci of impact” of Public Research Centres. The Marine Biology case helps us do this, because many of the research centres are located on the coast in peripheral regions and often⁴ far away from metropolitan areas. We see that for “Scientific Excellence”, the location of the research centres in peripheral regions plays less of roll, except that certain types of marine facilities are possible when located on the coast. However, our data so far indicates that, when looking at economic ties in the periphery, regional ties play a strong role.

This echoes a broader trend observed in urban studies and regions. Since the early 2000s, the literature has pointed to a trend towards agglomeration of R&D (production of knowledge) in large metropolitan areas where an increasing concentration of activities in these cities could be observed at the expense of others (Varga 2000, Duranton and Puga 2004, Fujita and Thisse 2013)). In this stream of literature, agglomeration is an important factor for R&D productivity in “Edison-type” application driven research whereas agglomeration is less important for “Pasteur-type” science driven research (Varga et al. 2014). In the latter type, *inter*-regional networks are emphasised more than *intra*-regional agglomeration. Recently a trend has been observed where the largest cities are undergoing a relative decline in a countries collective scientific activities (Grossetti 2013) with research moving away from the large cities elsewhere.

Our early findings leads us to see that collaborations in science in SBR follows the “Pasteur-type”, where there is little difference between regional linkages, national and international linkages. However, for the economic activities, the case shows “Edison-type” characteristics where regional linkages dominate. Our data so far also suggests that, the emergence of such regional districts does not happen on their own, they require institutional entrepreneurship (Garud et al. 2002, Robinson et al, 2007) through the forging of ties in the region through economic, regional policy and jobs (through professional training).

Although at early stages, this Research in Progress, speaks to the “strong notion” of peripherys and the importance of exploring this theme in terms of indicator development as well as for regional research and innovation policy more generally, for example in tailoring and assessing smart specialisation strategies (Foray et al. 2009).

⁴ Though not always, for example there is a large amount fo Marine Biotechnology research in Naples, Italy.

REFERENCES

- Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. *Handbook of regional and urban economics*, 4, 2063-2117.
- Foray, D., David, P. A., & Hall, B. (2009). Smart specialisation—the concept. *Knowledge economists policy brief*, 9(85), 100.
- Fujita, M., & Thisse, J. F. (2013). *Economics of agglomeration: cities, industrial location, and globalization*. Cambridge university press.
- Garud, R., Jain, S., & Kumaraswamy, A. (2002). Institutional entrepreneurship in the sponsorship of common technological standards: The case of Sun Microsystems and Java. *Academy of management journal*, 45(1), 196-214.
- Larédo, P., & Mustar, P. (2000). Laboratory activity profiles: An exploratory approach. *Scientometrics*, 47(3), 515-539.
- Malecki, E. J. (1980). Dimensions of R&D location in the United States. *Research Policy*, 9(1), 2-22.
- Marris, C., Joly, P. B., & Rip, A. (2008). Interactive Technology Assessment in the Real World Dual Dynamics in an iTA Exercise on Genetically Modified Vines. *Science, technology & human values*, 33(1), 77-100.
- Ritchie, R. J., Guy, K., & Philp, J. C. (2013). Policy to support marine biotechnology-based solutions to global challenges. *Trends in biotechnology*, 31(3), 128-131.
- Robinson, D. K. R., Rip, A., & Mangematin, V. (2007). Technological agglomeration and the emergence of clusters and networks in nanotechnology. *Research policy*, 36(6), 871-879.
- Varga, A. (2000). Local academic knowledge transfers and the concentration of economic activity. *Journal of Regional Science*, 40(2), 289-309.
- Varga, A., Pontikakis, D., & Chorafakis, G. (2014). Metropolitan Edison and cosmopolitan Pasteur? Agglomeration and interregional research network effects on European R&D productivity. *Journal of Economic Geography*, 14(2), 229-263.



Examining to what extent the source of competitive funding influences scientific impact. A case study of Danish EU FP7 funded projects.

Thomas Kjeldager Ryan^{*} and Jesper W. Schneider^{**}

^{*}tryan@ps.au.dk

Danish Centre for Studies in Research & Research Policy, Department of Political Science, Aarhus University, Bartholins Allé 7, Aarhus, DK-8000 C (Denmark)

^{**}jws@ps.au.dk

Danish Centre for Studies in Research & Research Policy, Department of Political Science, Aarhus University, Bartholins Allé 7, Aarhus, DK-8000 C (Denmark)

INTRODUCTION

To many researchers the ability to secure external competitive funding is a basic premise for engaging in substantial research activities and subsequent career progression. In many countries competitive funding is increasingly incentivised, but like other social activities in science, funding success is skewed with marked cumulative advantages to those who already have (Merton, 1968). The ability to obtain competitive funding is clearly connected with prestige, nevertheless, while “money is money”, funding instruments supporting “curiosity-driven”, “blue sky” research¹ may still be more appealing and prestigious to many researchers. Such funding instruments are typically less restrictive in their calls on matters such as research topics, collaborative requirements and societal impact. The latter requirements are often present in more strategic funding programmes and especially the European Framework Programmes have stipulated a social contract between science and society in Europe, increasingly requiring research to be oriented towards addressing social, economic, cultural, and political challenges. Application for such “challenged-oriented” international research funding is perceived by some as over-complicated and too resource demanding and anecdotal evidence from Denmark suggests that high performing research groups opt to bypass such calls because they prefer and indeed are able to secure sufficient “curiosity-driven” funding with less obligations (UFM Report, 2015).

While the perception of prestige is difficult to challenge, the often held assumption that research funded by “curiosity-driven” instruments in general lead to higher academic impact measured by citations compared to “challenge-oriented” research, with its supposed focus on societal relevance and applicability, can indeed be examined. A recent study examining the impact of Danish publications linked to the European Framework Programmes suggests that there may be no contradiction between research being “challenge-oriented” and being academically influential when it comes to citation impact (Schneider & Ryan, 2015; Ryan, Schneider & Mejlggaard, 2015).

¹ “[blue sky research] is scientific research in domains where “real-world” applications are not immediately apparent. It has been defined as research without a clear goal and “curiosity-driven science”
https://en.wikipedia.org/wiki/Blue_skies_research.

In one of these analyses, we examine the citation impact of three sets of journal articles published between 2005 and 2011 (Schneider & Ryan, 2015). The three sets contain Danish publications linked to two national “curiosity-driven” funding instruments, and one international “challenge-oriented” programmes, i.e. the European Framework Programme FP7. The two national “curiosity-driven” funding instruments are prestigious individual career grants from the Danish Research Council for Independent Research (DFF), typically running for 2 to 4 years; and the exclusive and very prestigious Centre of Excellence instrument funded by the Danish National Research Foundation (DNRF), typically running for 10 years. These are the two most prestigious national awarding bodies in Denmark explicitly supporting basic, curiosity-driven “blue sky” research. Previous findings have confirmed that the DNRF and DFF publication sets have high impact but also that the impact of the DNRF set is highest (Schneider & Costas, forthcoming). Both findings were expected because both instruments were conceived of as supporting high quality applications, and the different modes and durations of the instruments suggests that the excellence initiative in general would have slightly higher impact.

The surprising findings from the recent study suggest that the performance of papers linked to the “challenge-oriented” European Framework Programmes is on par or slightly above those of the prestigious Danish “curiosity-driven” instruments. The results are challenging not least because the average citation impact of Danish papers is among the highest in the world². The results therefore need to be scrutinized further in order to establish the factors behind the seemingly very high impact of “challenged-oriented” research papers and this is the aim of the present work. We specifically examine to what extent the source of competitive funding matters for scientific impact in a Danish context.

Previous research on the impact of funding sources is both sparse and inconclusive. Early work on links between impact and the funding body was pioneered by e.g. Lewison (1994) and Lewison and Dawson (1998). Within the evaluation literature, studies generally focus upon potential impact effects of being funded by a certain funding body versus not being funded by that particular body (for a critical overview, see Schneider & van Leeuwen, 2014). The European Framework programmes and the European Research Council have been subjected to various investigations but scholars tend to examine other aspects than impact (e.g., Luukkonen, 2012; Neufeld, Huber, & Wegner, 2013). Most recently, Schneider and Costas (forthcoming) have examined the presumed impact differences in publication sets belonging to two different public Danish funding bodies. Part of this data set is included in the present study. Finally, Gök, Rigby and Shapira (2016) utilize the funding acknowledgements in Web of Science (WoS) to examine the potential relationships between impact and sources of funding.

The rest of the paper is organized as follows, data and methods are presented in the next section; the subsequent section presents some of the most important results; and we end with a general discussion of the findings.

²See this recent report: http://ufm.dk/publikationer/2015/filer/dfir_scientometric_analysis_final.pdf

METHODS

The dataset in this paper is a compilation of three separate publication sets: Danish publications linked to FP7, DFF and DNRF. The three publication sets are unique because they are validated by Danish grant recipients as being an output of their FP7, DFF and/or DNRF funding. As mentioned above, these three funding instruments are very different and can be considered to represent national “curiosity-driven” funding (i.e. DFF and DNRF) and international “challenge-oriented” funding (i.e. FP7). Together the three validated publication sets constitute an adequate basis for a modelling approach in order to contrast and scrutinize the factors associated with the underlying citation patterns (e.g., Didegah & Thelwall, 2013).

Together, the three publication sets contain 11,841 journal articles published between 2007-2011 and indexed in WoS. The FP7 set contains 1,908 articles, the DNRF set 5,582 and the DFF set 5,774. Finally, of the 11,841 articles, 1,353 are linked to two or more of the funding programmes. We utilize the enhanced WoS-database at CWTS, Leiden University to obtain publication data as well as field normalized citation indicators for articles and journals.

First, we examine the differences in impact between the publication sets using the Mean Normalized Citation Score (MNCS) with a two-year citation window. We also describe the type of research linked to each funding instrument. These preliminary results inform the second part of the analysis where we apply Ordinary Least Squares (OLS) regression to estimate differences in impact while attempting to statistically “control” for initial differences between the funding sources and other potential confounding factors.

Modelling approach

We estimate differences in impact by running OLS models where articles’ continuous citation scores are transformed to fit a continuous lognormal distribution. One is added to the individual articles’ field normalised citation scores (NCS), then the natural logarithm to the score plus one is taken. Subsequently we estimate models with multiple independent variables using OLS regression. This is in sync with the recommendations by Thelwall and Wilson (2014). They find that citation data fit a discrete lognormal distribution but in our case due to the field normalization, data are continuous, which can only improve the fit of the model as the dependent variable is continuous by definition and thus likely follows a continuous lognormal distribution.

The independent variables of interest in the regression include FP7 and DNRF which are dummy variables indicating whether the articles are linked to the particular funding source. Funding from DFF is not directly included in the analysis and is therefore the reference category. While this modelling approach allows us to improve the fit of the model and still includes the uncited articles, it makes a direct interpretation of coefficients somewhat complicated.

The different funding instruments and bodies fund research that is different on many parameters. Our modelling approach therefore tries to level-out these differences because we want to be able to estimate, given that the funded research is alike, to what extent there is a difference in impact between “curiosity-driven” and “challenge-driven” research.

In the regression specifications we therefore “control” for factors that are known to correlate with citation impact and that may confound with the estimated impact of the funding source. Theoretically we can thereby divide the difference in impact between funding sources as something that can be attributed to known differences in the type of research that is funded

and to the type of funding itself. Obviously there will be residual differences which we cannot explain.

The “control” variables include: 1) countries included in the address field; 2) the research field in which the article is published; 3) logged number of authors; 4) the number of self-citations; 5) the average number of self-citations per author; 6) the number of references; 7) the number of covered references in WoS; 8) the normalized journal citation score (NJS); 9) whether or not the article had more than one collaborator from another institution; 10) whether there were any industrial addresses; 11) whether the article was a long-distance collaboration; and 12) number of countries. Notice, by using field normalised citation scores we have in principle made comparisons across fields possible without the need to “control” for this in the specification. The basic model specification is therefore:

$$\text{Log}(\text{NCS}+1) = \beta_0 + \beta_1 \text{FP7} + \beta_2 \text{DNRF} + X\beta_{\text{controls}} + \varepsilon$$

Where $X\beta_{\text{controls}}$ is the vector of control variables and ε is the disturbances.

Time dimension of impact

Notice, it is possible that an initial difference in impact between articles based on the source of funding may decline over time because especially European Framework Programmes fund research where an international (European) set-up is obligatory and that this may give a head start in terms of citation impact compared to funded research without such requirements (albeit the latter can certainly also be international to some degree). We therefore run the standard specification for five different citation windows (1-year to 5-year citation windows). This allows us to investigate whether impact differences are temporal or whether there is a sustained difference between the impact of research output depending on the funding source. Thus the regression equation above is repeated for $t = \{1, 2, 3, 4, 5\}$, where t is the citation window.

RESULTS

Table 1 below presents descriptive statistics for MNCS with a two-year citation window. These are the basic comparisons that triggered the main research question. Overall the impact is high for all funding programmes compared to the national average (1.47) (Schneider & Ryan, 2015). Impact is highest for FP7 followed by DNRF and DFF. However, if we remove the overlap in which publications are funded by two or three of the sources then the impact values level out and a clear difference is hard to discern. The articles with the overlap have a very high impact, with the mean larger than the 75th percentile indicating a very skewed distribution that is pushed to the right in general compared to the non-overlap distributions.

Table 1 Mena Normalized citation score two-year citation window 2007-2011.

With overlap								
	Mean	Std dev	1 Quartile	Median	3 Quartile	Min	Max	No. pubs
FP7	2.16	5.39	0.25	0.97	2.36	0.00	123.39	1908
DNRF	1.94	3.37	0.22	0.95	2.27	0.00	45.36	5582
DFF	1.85	4.10	0.00	0.85	2.05	0.00	123.39	5774
Without overlap								
	Mean	Std dev	1 Quartile	Median	3 Quartile	Min	Max	No. pubs
FP7	1.85	4.28	0.00	0.91	2.12	0.00	106.72	1431
DNRF	1.78	3.06	0.00	0.87	2.14	0.00	43.86	4403
DFF	1.64	3.52	0.00	0.78	1.84	0.00	93.08	4670
The overlap								
	Mean	Std	1 Quartile	Median	3 Quartile	Min	Max	No. pubs
DFF DNRF & FP7	3.07	7.76	0.43	1.14	2.87	0.00	123.39	477
DFF & DNRF not FP7	2.51	4.17	0.30	1.28	2.82	0	45.36	876

This is however, not enough to establish whether impact is higher for one group or another because at this moment we are comparing articles written within a variety of fields, years, in collaboration with different countries and in different context. We want to estimate whether there is a difference in impact for the same type of research but with different types of funding. Therefore, the next task is to segregate impact differences into types of research and types of funding. Below we look at how research differs depending on the funding source.

While the three funding instruments overlap in terms of the research fields they fund, they differ in terms of the proportions to which they do so. First of all, articles linked to FP7 are more frequently within the engineering and agricultural sciences relative to DFF and DNRF, while articles linked to DFF are more frequently within biology, biomedicine and physics relative to FP7. Lastly, DNRF has a relatively high share of articles within chemistry and physics. DFF and DNRF have a wide scope of collaborating countries but are highly concentrated within USA. DNRF has the highest proportion of collaborating authors in USA, Canada, Australia China and Japan. While DFF differs from FP7 in relation to countries only in terms of how much they collaborate with UK and USA, for the other countries it is approximately the same distribution.

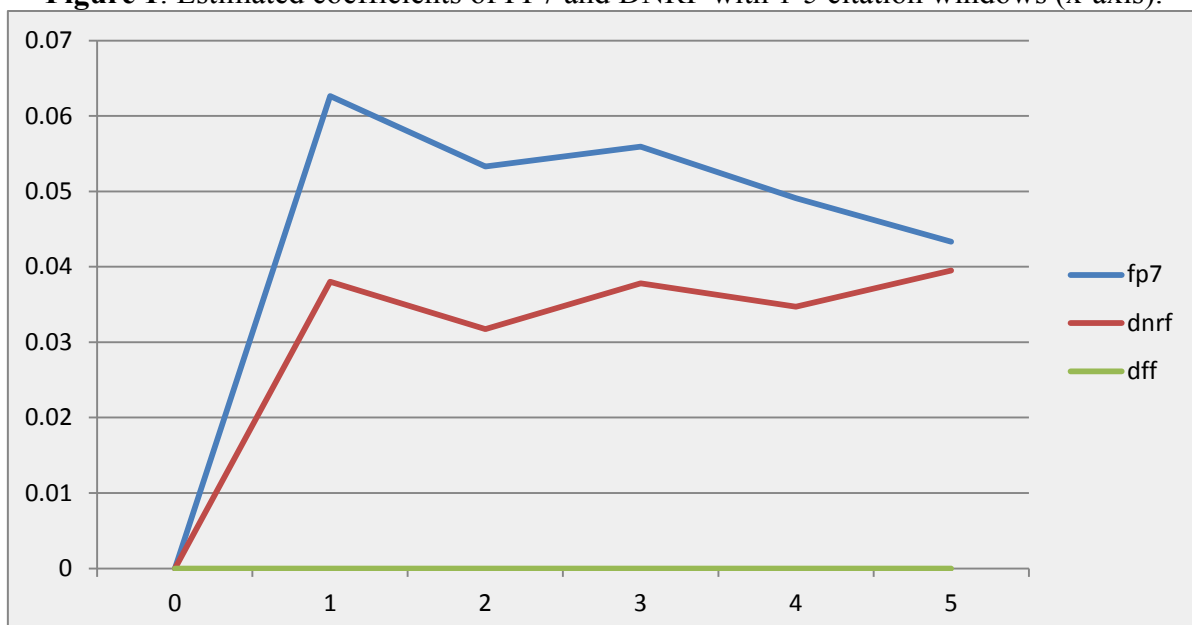
In general, FP7 has the largest fraction of articles including authors from foreign institutions followed closely by DNRF. The average number of authors, institutions and countries is highest for DNRF with a substantial gap down to FP7 and DFF. Articles funded by FP7 are twice as likely to have industrial authors as those funded by DFF and DNRF. Articles funded by FP7 and DFF are published in journals with a normalized journal citation score 0.1 points below those of DNRF on average. When articles are funded by both FP7 and one or two of the national funding instruments, articles are published within the basic fields of science and authors from institutions in the UK, USA and Germany are the most common collaborators. In terms of coverage FP7 articles have a lower coverage (i.e. proportion of references covered

by WoS) compared to DFF and DNRF. This is most likely because of the difference in subject profiles, i.e. engineering for example generally have lower coverage in WoS.

The takeaway message is thus, FP7 is linked to more European research mainly within the technical and engineering sciences and have a high industrial involvement. The coverage is lower which may be related to the type of research on average being more applied (i.e. conference papers are used more often within technical sciences and these are covered less than journal articles in basic sciences). The DNRF publication set is the most international and the DNRF funds larger projects with many authors and countries. DFF has a lower impact in general (though still high compared to the national average) and fund relatively smaller projects.

We now turn to the regression results. The basic results are shown in Table 2 in the Appendix.

Figure 1: Estimated coefficients of FP7 and DNRF with 1-5 citation windows (x-axis).



Note: The figure shows plots the coefficients from the 5 regressions in the Appendix. DFF is the reference group, why it is zero for every regression. The x-axis indicates the citation window while the y-axis is the coefficient of FP7 or DNRF in terms of $\text{Log}(\text{nCS}+1)$. The interpretation is thus, the effect of a particular funding programme on impact relative to the base-case (DFF).

The regression results show that when statistically “controlling” for the initial presumed differences, there is a difference in the expected impact (NCS+1) of 6 percentage points compared to DFF and 2 percentage points to DNRF. After 5 years the initial difference between FP7 and DNRF has nearly disappeared and the difference between DFF and FP7 has decreased to 4 percentage points.

Interestingly, it seems that the differences observed in the initial study using 2- or 3-year citation windows are not a lasting impact difference (Schneider & Ryan, 2015).

We discuss what could be the reason for this temporal difference in the final discussion section.

In conclusion, our study finds that the funding source does correlate with impact and “challenged-driven” FP7 funding have a marginally higher impact compared to more “curiosity-driven” national funding instruments, but most importantly the set of articles linked

to FP7 has a higher impact in quite different areas of research with somewhat different international collaborators, thus, the funding instruments clearly complement each other. We should caution that the main findings are valid for the Danish context only. Generalizations beyond this context are presently unfounded. The data set should be enlarged with other countries and funding sources, or complimentary and comparable national case studies need to be carried out in order to assess the external validity of the present findings.

DISCUSSION

In this paper we asked to what extent different funding sources are related to citation impact. Using three large validated publication sets of articles funded by two Danish and one European funding organization we attempted to assess the differences in impact between two conceptually different forms of funding: “challenged-driven” funding represented by FP7 and more “curiosity-driven” funding represented by two different Danish funding instruments, an excellence initiative (DNRF) and a more individual career initiative (DFF). A previous study showed us that the overall impact for articles linked to FP7 funded research was higher than both the DNRF and DFF on average when measured by MNCS with a two-year citation window. However, the research funded under each programme is very different in terms of field of research and number and choice of collaborative partners. In addition, the very difference in how and which research is funded may also affect how fast citations are accumulated, as well as the profile and half-life of the ensuing citation distribution of such articles, hence impact differences between the publication sets may initially be high in the beginning but after a few years they may converge or even reverse.

Using OLS regression we find that there is a difference in expected impact dependent on funding, however the impact differential of funding decreases over time. Thus, indicating that FP7 “challenge-oriented” funding is cited faster but not necessarily more than “curiosity-driven” nationally funded publication sets. We also find that articles linked to DFF-funding (i.e. shorter running individual grants) have a lower impact over time, while publications linked to the generally larger DNRF and FP7 grants end up having approximately the same impact over time when controlling for other differences.

While we have been talking about impact differences in this paper it is noteworthy that all funding programmes have MNCS scores far above the national average. Also that they have very different target groups. Thus, it may be more relevant to see the three funding regimes as three complementary high impact funding regimes that not only fund different types of research, but also seemingly different research fields and support research collaborations in various regions.

APPENDIX

Table 2: Basic regression results

	Model t=1		Model t=2		Model t=3		Model t=4		Model t=5	
Number of obs	11,841		11,841		11,841		11,841		11,841	
F(46, 11794)	25.82		54.05		61.89		59.84		58.38	
Prob > F	0		0		0		0		0	
R-squared	0.19		0.32		0.37		0.39		0.40	
Root MSE	0.78		0.57		0.50		0.47		0.46	
	Robust		Robust		Robust		Robust		Robust	
Log(ncs+1)	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
	t=1		t=2		t=3		t=4		t=5	
FP7	0.063	0.021	0.053	0.016	0.056	0.014	0.049	0.013	0.043	0.013
DNRF	0.038	0.015	0.032	0.011	0.038	0.010	0.035	0.010	0.039	0.009
Self citations	0.075	0.009	0.030	0.004	0.019	0.003	0.015	0.002	0.013	0.002
self citations per Author	0.629	0.077	0.173	0.023	0.117	0.013	0.086	0.010	0.073	0.009
NJS	0.125	0.008	0.155	0.006	0.152	0.006	0.145	0.007	0.140	0.007
References	0.000	0.001	0.001	0.001	0.002	0.001	0.002	0.001	0.002	0.001
Covered References	0.003	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.000	0.001
Colaboration	0.001	0.021	0.026	0.015	0.020	0.013	0.012	0.013	0.007	0.013
Industry	0.014	0.028	0.002	0.020	0.000	0.018	0.005	0.017	0.005	0.017
Long distance Collabration	0.019	0.020	0.007	0.015	0.000	0.013	0.000	0.012	0.006	0.012
Log(Authors)	0.066	0.015	0.080	0.013	0.076	0.013	0.067	0.013	0.067	0.014
Number of Contries	0.001	0.006	0.004	0.005	0.008	0.005	0.010	0.005	0.012	0.005
Country Dummies	Yes		Yes		Yes		Yes		Yes	
Research Area dummies	Yes		Yes		Yes		Yes		Yes	

REFERENCES

- Didegah, F., & Thelwall, M. (2013). Determinants of research citation impact in nanoscience and nanotechnology. *Journal of the American Society for Information Science and Technology*, 64(5), 1055–1064.
- Gök, A., Rigby, J., & Shapira, P. (2016). The impact of research funding on scientific outputs: Evidence from six smaller European countries. *Journal of the Association for Information Science and Technology*, 67(3), 715–730.
- Lewis, G. (1994). Publications from the European-Community Biotechnology Action Program (BAP) - Multinationality, acknowledgment of support, and citations. *Scientometrics*, 31(2), 125–142.
- Lewis, G., & Dawson, G. (1998). The effect of funding on the outputs of biomedical research. *Scientometrics*, 41(1–2), 17–27.
- Luukkonen, T. (2012). Conservatism and risk-taking in peer review: Emerging ERC practices. *Research Evaluation*, 21(1), 48–60.
- Merton, R. K. (1968). The Matthew effect in science. *Science*, 159(3810), 56-63.

- Neufeld, J., Huber, N., & Wegner, A. (2013). Peer review-based selection decisions in individual research funding, applicants' publication strategies and performance: The case of the ERC Starting Grants. *Research Evaluation*, 22(4), 237–247.
- Ryan, T.K., Schneider, J.W. & Mejlgaard, N. (2015). Can society-oriented research be academically influential? *Euroscientists*. <http://www.euroscientist.com/can-society-oriented-research-be-academically-influential/>
- Schneider, J.W. & Costas, R. forthcoming. Identifying potential breakthrough publications using refined citation analyses: Three related explorative approaches. Accepted for publication in: *Journal of the Association for Information Science and Technology* – available at: <http://arxiv.org/abs/1512.01388>
- Schneider, J.W. & Ryan, T. K. (2015). Bibliometric analyses of publications from Danish PIs linked to EU Framework Programs 6 and 7. http://ufm.dk/publikationer/2015/filer/technical-annex_1_bibliometric-analyses-of-publications-from-danish-pis-linked-to-fp.pdf
- Schneider, J. W., & van Leeuwen, T. N. (2014). Analysing robustness and uncertainty levels of bibliometric performance statistics supporting science policy. A case study evaluating Danish postdoctoral funding. *Research Evaluation*, 23(4), 285-297.
- Thelwall, M., & Wilson, P. (2014). Regression for citation data: An evaluation of different methods. *Journal of Informetrics*, 8(4), 963-971.
- UFM. Ministry of Higher Education and Science (2015) Effects of participation in EU framework programmes for research and innovation – for researchers, institutions and private companies in Denmark. <http://ufm.dk/en/publications/2015/effects-of-participation-in-eu-framework-programmes-for-research-and-innovation-for-researchers-institutions-and-private-companies-in-denmark>



The Determinants of National Funding in Trans-national Joint Research: Exploring the Proximity Dimensions¹

Emanuela Reale^{*}, Andrea Orazio Spinello^{**} and Antonio Zinilli^{***}

^{*}emanuela.reale@ircres.cnr.it

Research Institute on Sustainable Economic Growth (IRCRES), National Research Council (CNR), Via dei Taurini 19, Rome, 00185 (Italy)

^{**}andrea.spinello@ircres.cnr.it;

Research Institute on Sustainable Economic Growth (IRCRES), National Research Council (CNR), Via dei Taurini 19, Rome, 00185 (Italy)

^{***}antonio.zinilli@ircres.cnr.it

Research Institute on Sustainable Economic Growth (IRCRES), National Research Council (CNR), Via dei Taurini 19, Rome, 00185 (Italy)

INTRODUCTION

This paper investigates -using an explorative approach, why policy makers at national level engage in transnational joint research activities and mobilize dedicated financial resources. The research question is: why policy makers (either Governments or Research Funding Organisations-RFOs) in EU28 countries invest in transnational joint research activities beyond the European Framework Programmes, and what are the determinants of different levels of funding engagement? The question is relevant to understand the reasons that generate the existing imbalances within European countries as to the participation in transnational research, which are likely to create peripheries within the ERA, thus undermining the process of European integration.

We assume that proximity linked to cognitive, institutional and organizational dimensions can affect the policy decisions about the level of funding (real engagement) joint European research programmes, because the closeness or distance in these dimensions generate similarities that are likely to influence the possibility of decision makers to collaborate in the implementation of research programmes. The paper also explores the existence of any effect of geographical proximity, although it is not supposed to play a role in policy decisions about investment in transnational research programmes.

THE CONCEPTUAL FRAMEWORK

Boschma (2005) defines proximity an “umbrella concept” including several dimensions - physical, cognitive, social, organisational and institutional, which represent the bases for interactions between agents. The dimensions can overlap, their importance can change over time and they might produce both positive and negative effects.

The concept of proximity was generally applied at the performer level, in inter-firm collaborations, innovation and regional economic development (Arnard and Khanna, 2000; Boschma, 2005) assuming that the geographical closeness is likely to produce certain types of effects facilitating collaborative patterns. Several empirical investigations have been produced, and a bulk of literature exists, where the importance of geographical proximity is

¹ The names of the authors are listed in alphabetic order. This work was supported by RISIS Research infrastructures Project under the EUFP7 Programme

generally confirmed (Arundel and Geuna, 2004; Laursen et al., 2011), and other dimensions of proximity emerged as well (e.g. cognitive proximity as to the absorptive capacity of firms to collaborate with the universities, Dreier et al., 2014). The different dimensions of proximity are generally used in order to benchmark their prominence with respect to the geographical one, and to understand the way in which the different dimensions interact producing effects (Frame and Carpenter, 1979).

Frenken et al. (2009) review the literature on spatial scientometrics and the role of proximity in collaborative innovation, using the five dimensions proposed by Boschma (2005), in order to propose a model combining spatial scientometric papers and proximity dimensions to investigate scientific collaborations. Knoblen and Oerlemans, (2006) as well present a review of the way in which the different dimensions of proximity are defined in the literature putting into evidence inconsistencies and overlapping in the content that produce some uncertainties in the interpretation of the results.

In this work we explore the possibility to use non-spatial proximity and tentatively the geographical proximity to understand how being close or distant in the quoted dimension affects (positively or negatively) the decision makers (governments and RFOs) at national level to engage in participation and mobilize funding of transnational joint research programmes. We argue therefore, that the decisions about the level of funding engagement in the transnational joint programmes investigated, are influenced by the presence of factors that tend to favour processes of coordination between countries because of the closeness between decision makers at national level.

Non-spatial proximity includes several dimensions, representing the extent to which the same research policy structure and hierarchical control over the decisions characterize the countries where the agents operate, and the excellence of the research base. *Institutional* proximity identifies the extent to which decision makers in different countries have similar levels of investment in R&D, have the same level of government effectiveness, and the sharing of research priorities. *Cultural* proximity identifies the closeness in the language and cultural background between countries, including administrative traditions; *cognitive* proximity identified the presence of a shared knowledge base and interactions between the local research communities, because of the high standing quality of the knowledge production. *Organizational* proximity identifies the extent to which decision makers in the different countries are under similar structures of hierarchical control, and own similar coordination power as to R&D policy (Rallet and Torre, 1999; Knoblen and Oerlemans, 2006). Finally, *Geographical* proximity is the physical distance between decision makers (RFOs) in charge for formal and real engagement in transnational joint research programmes. The kilometric distance between the main towns of the countries where the RFOs are localized operationalizes the distance.

METHODOLOGY

The effect of proximity on the level of countries engagement in joint transnational research programmes, is investigated through the JoREP database. JoREP stores descriptive information on the characteristics of research programmes, RFOs managing the different programmes, and on the volume of funding channelled through these programmes. It allows analyses of the modes of the Europeanization, figuring out the national strategies of participation and funding mobilization. To study the transnational joint research programmes we focus on the participation and the funding in two type of coordinated programmes: ERA-NETs and JPIs in EU28 plus four associated countries, from 2010 to 2014 (five years), using data collected in the JoREP dataset. The dataset includes 47 ERA-NETs and 9 JPIs that launched at least one call in 2013 or 2014.

The variables included in the dataset and used for the analysis are:

- **Agency country.** The country where the funding agency is established
- **Geographical level of the funding agency.** Funding agencies are distinguished between national and regional. This distinction refers to the institutional embedding, not to the funding activities; e.g. a regional agency, funded under regional law, might support also research outside the region.
- **Agency classification/agency policy domain.** The classification of funding agencies is a two-level one; the first level refers to the position with respect to the State (Government, Agency, and Performer) while the second one specifies more precisely the policy domain of activity (Science Ministry, Sectoral Ministry, Regional Government, Research Council Innovation Agency, Sectoral Agency, Public Research Organization, Private research organization).
- **Research topics.** For classification of programme topics, the Nomenclature for the Analysis and Comparison of Scientific Programmes and Budgets (NABS) from the Frascati Manual (2007 version) is adopted. This classification refers to the socio-economic objective of the programme, not to the actual research content.
- **Type of programme.** This descriptor defines the typology of the research programme (ERA-NET, ERA-NET PLUS, JPI).
- **Amount.** The earmarked funding for the whole year expressed in Euro for each year of reference, as reported in the programme call or in official documents/websites.
- **GERD as % of GDP** (source: EUROSTAT).
- **Funding by NABS** – R&D Policy objectives that approximate the national priorities.
- **Researchers on population.** The ratio between total number of researchers and total population for each country considered (source: EUROSTAT).
- **H-Index.** Number of articles produced in a country (h) that have received at least h citations (Source SCOPUS).
- **Patents on population.** The ratio between total patents and total population (resident and non-resident) for each country considered (source: OECD).
- **Language.** Official language in the country.
- **National Administrative Tradition.** The variables classify the countries on the base of the administrative tradition according to classification proposed in the literature (Pollit and Bouckaert, 2004). Four categories of traditions are identified: Napoleonic, Anglo-Saxon, Germanic, Scandinavian-Nordic, which correspond to different countries.
- **Government Effectiveness.** It represents the capability of a country to effectively formulate and implement sound policies (source: World Bank, Kaufmann et al., 2010).
- **Currency.** National currency of the country.

The statistical approach used to investigate proximity is a spatial panel model to explain the dependent variable “amount of funding”. The spatial units are the countries that fund the calls within programmes, and the analysis deals with the time effects from 2010 to 2014. The panel data are generally more informative, indeed we can identify time effects able to bring out latent processes, as well as the positive effects of a purely statistical (for instance the reduction of collinearity among the variables) and hence increases efficiency in the estimation. In the context of spatial analysis, the relevant dimensions are shown in the prospect below.

Non-spatial proximity	Institutional proximity	GERD as % of GDP. Researchers on population. Entropy of policy objectives (NABS). Government Effectiveness.
	Cultural proximity	The country language. National Administrative Traditions.
	Cognitive proximity	H-Index as a proxy of the closeness in the quality of scientific performance in the countries. Patents as a proxy of the technological orientation of the country.
	Organizational proximity	The Agency classification with respect to the state The RFOs policy domain of activity.
Spatial proximity	Geographical proximity	The weight matrices is based on the centroid distances, d_{ij} , between each pair of spatial units i and j .

Entropy is a measure of heterogeneity, which shows the diversity of the socio-economic objectives by countries over the years. Entropy is 0 when all Countries are in the same socio-economic objective and it is maximal when the number of different objectives increases. Knowing that the presence of spatial interdependence is manifested by phenomena of spatial concentration of similar values (in the case of positive interdependence) or of different values (in the case of negative interdependence) we use the Moran's index that represents the more traditional measure of spatial correlation.

Two models have been built, one for each type of programme. In the model the dependent variable is given by the logarithm of the amount of funding. Independent variables used in the models are those listed in the prospect plus Currency.

DATA ANALYSIS

Descriptive analysis

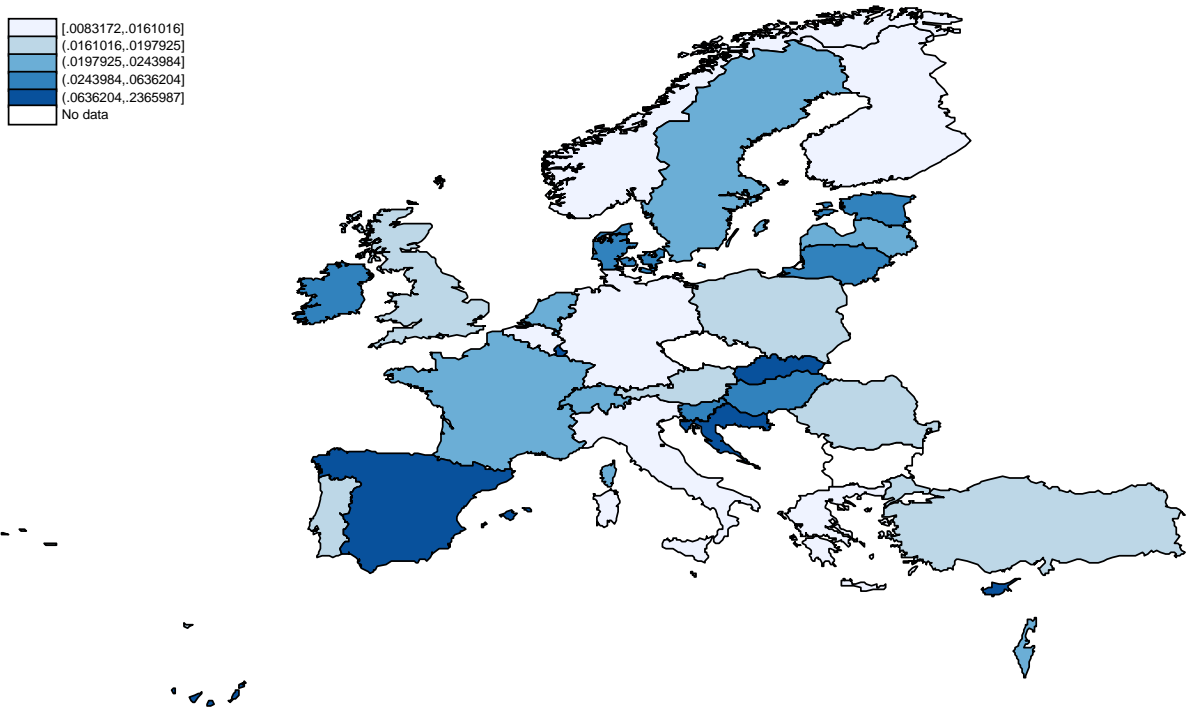
Table 1 shows some descriptive statistics about the amount of funding in the two types of programmes from 2010 to 2014. First of all we notice the differences on the share of funding between the two types of programme, the volume of funding mobilized in ERA-NETs is higher than JPIs. The results illustrate that the values seem to be constant over the years, whereas for JPI programmes we observe a large increase of standard deviation in 2013.

Tab. 1 - Descriptive Statistics on amount over the years by type of programme (KEuros)

	Amount-ERA-NET					Amount-JPI			
	2010	2011	2012	2013	2014	2011	2012	2013	2014
Mean	656.473	843.234	853.893	919.956	838.053	711.363	762.5	846.172	627.931
Std. Dev.	717.819	863.135	1,013.077	969.127	888.334	483.767	438.035	916.219	562.706
Min	20.000	21.000	38.000	10.000	20.000		200.000	150.000	100.000
Max	4,000.000	5,000.000	7,300.000	8,383.430	7,200.000	2,000.000	1,500.000	6,000.000	3,000.000

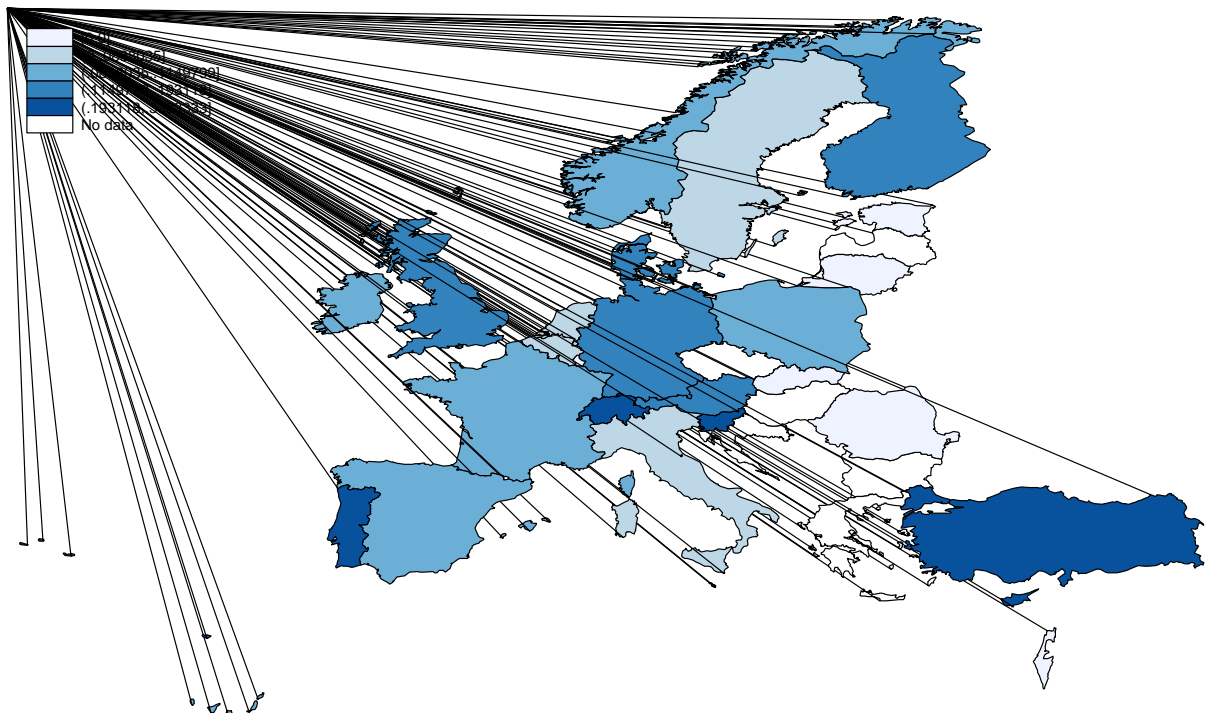
Figure 3 shows the entropy distribution among countries in the ERA-NET programmes (the entropy is standardized on the base of participation).

Fig. 3 - Spatial distribution by NABS in the ERA-NET programmes



Slovakia, Cyprus and Croatia have a greater heterogeneity in terms of NABS about the programmes, followed by Lithuania, Spain and Luxembourg. Below (Fig. 4) the entropy distribution among countries in the JPI programmes (the entropy is standardized on the base of participation).

Fig. 4 - Spatial distribution by NABS in the JPI programmes



In the JPI programmes we can observe another type of configuration, indeed in this case the countries with the highest degree of heterogeneity are Slovenia, Portugal, Cyprus and Switzerland. Germany, Finland, Austria and Turkey belong to another cluster. Interestingly enough, most of the countries with very low entropy are all located in East Europe.

In Table 2 we have calculated the Pearson's correlation between H-index and amount of funding by year for each programme.

Tab. 2 - Pearson's correlation between H-index and amount of funding

Type of programme	2010	2011	2012	2013	2014
ERA-NET	0.4090***	0.4844***	0.4205***	0.4153***	0.3662***
JPI		0.3356	0.0890	0.3935***	0.3486***

Significant: ***0.01, **0.05, * 0.10

In the correlations above we notice that there is a relationship between the two variables, although if it seems to decrease in the last two years considered.

Spatial correlation

We proceed to the calculation the index of Moran on the variable amount of funding for ERA-NETs and JPIs obtaining the following results (tab. 3):

Tab. 3 - Moran's Index on amount of funding by year

	2010	2011	2012	2013	2014
ERA-NET	-0.188***	-0.140***	-0.125***	-0.95***	-0.059***
JPI		-0.368***	-0.348**	-0.246***	-0.141**

Significant: ***0.01, **0.05, * 0.10

The results show a negative spatial correlation over the years; negative spatial autocorrelation is when dissimilar values are close. During the years we can observe decreasing trends of the Moran's coefficient, thus there is a convergence in the amount allocated (for both programmes).

Estimation results

Table 4 shows the estimates of final spatial panel model with random effects used to study the amount of funding for JPI and ERA-NET programmes.

Table 4 – Estimates

Variable	Eranet			JPI		
	(1)	(2)	(3)	(1)	(2)	(3)
WLogamount	-.0009355***	-0.0006478***	-0.00102***	.0114***	-0.00324**	-0.00500***
H (Entropy of Nubs)	232.4295 ***			73.22518***		
CenteredGerd		.413*			0.615*	
Gov. Effectiveness		8.645***			6.550***	
Germanic tradition		-0.101			-1.418***	
Napoleonic tradition		0.436***			0.31	
Scandinavian tradition		-0.658***			-2.069***	
Anglo-Saxon tradition		-1.112***			-0.474	
H Index			0.0196***			0.0191***
Patents			-0.000138***			-0.000170***
Constant	3.810211*	1.406***	2.168***	-5276341	1.994***	1.536***

*** p<0.01, ** p<0.05, * p<0.1

The variable ‘spatial lag’ (WLogamount) able to measure the effect of the nearby countries of the dependent variable appears to be significant and has a negative sign in both programmes, which signal the existence of a spatial negative correlation. This means that neighbouring areas are different from those distant, in particular with increasing distance increases similar values of the dependent variable considered. The Entropy is significant (with a very high value), meaning that with increasing the number of socio-economic objectives it increases the share of the investment. Concerning GERD and Government Effectiveness effect we notice that they are significant and have a positive effect on the volume of funding for both the programmes. The H-index variable is also significant in the two programmes, this means that the H-index of the country influence the share of funding positively. Patents variable turns out to be significant but negative.

CONCLUSIONS

The paper investigates the level of engagement of different European countries in transnational joint research activities beyond the EUFPs. We test the institutional, cultural, organizational, and cognitive proximity as characteristics that can predict the level of funding a country is likely to mobilize, and the geographical proximity for effects linked to the localization of a given country in Europe.

The focus was on the EU28 plus four associated countries participation and funding from 2010 to 2014 in different transnational programmes, namely ERA-NETs, ERA-NET PLUS, and JPIs. Some interesting evidences emerged, no matter whether the programmes are ERA-NETs or JPIs, namely:

- As to the non-spatial proximity, three factors are likely to produce a growing funding mobilization: the closeness in scientific performance, the closeness in government effectiveness, and the closeness in national R&D expenditure (the higher they are the higher the funding mobilization in joint transnational research). Moreover the heterogeneity of political research objectives is positively associated with an increasing investment at national level. Thus, cognitive proximity and institutional proximity are more influential than organizational and cultural proximity on the policy maker’s decisions to invest in trans-national joint research activities.
- Surprisingly enough, the geographical proximity plays a role. Countries are likely to invest the same amount of money of other countries when the geographical distance is high. Saying differently, the neighbouring countries have different amounts of funding mobilization in transnational research. It means that despite the institutional and cognitive proximity and proximity in R&D expenditures, neighbouring countries have different levels of funding mobilization.

A sort of “similarity effect” emerges, which is shaped by the self-reinforcing effects produced by cognitive and institutional proximity. It explains the factors influencing high levels of research funding mobilization in some European countries and the low funding investment in others. Despite the existence of a general push toward global integration, closeness and distance between countries matters as far as the process of integration is concerned, which are likely to reinforce the existing imbalances, creating ‘peripheries’ within Europe.

REFERENCES

- Anand, B., Khanna, T. (2000) “Do Firms Learn to Create Value? The Case of Alliances”, *Strategic Management Journal*, 21: 295-315.
- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Boston: Kluwer Academic.
- Arundel A. and Geuna A. (2004), Proximity and the use of public science by innovative European firms. *Economics of Innovation and New Technology* 13(6) 559–580 13:6, 559-80.
- Boschma R. (2005) “Proximity and Innovation: A Critical Assessment”, *Regional Studies*, 39:1, 61-74, DOI: 10.1080/0034340052000320887
- Claeys P., Moreno R. and Suriñach J., 2010, Fiscal Policy and Interest Rates: The Role of Financial and Economic Integration, in: Paez, A., Le Gallo J., Buliung R., Dall’Erba S. (eds.), *Progress in Spatial Analysis: Methods and Applications*, Springer: Berlin, p. 311-338
- Fernández, Diestro A. y Valle López J. M. (2015). Towards a European Supranational Policy of Education based on the European Dimension on Education. *Bordón. Revista de Pedagogía*, 67(1), 101–116. doi: <http://dx.doi.org/10.13042/Bordon.2015.67107>
- Frenken, K. (2009): Proximity, Social Capital and the Simon Model of Stochastic Growth. In Reggiani, A. – Nijkamp, P. (eds): *Complexity and Spatial Networks. In Search of Simplicity*. Springer-Verlag, Berlin–Heidelberg, pp. 133-140.
- Hoekman, J., Frenken, K., & Tijssen, R. J. W. (2010) “Research collaboration at a distance: changing spatial patterns of research collaboration in Europe” *Research Policy*, 39, 662–673.
- Kaufmann D., Kraay A., Mastruzzi M. (2010). *The Worldwide Governance Indicators. Policy Research*, The World Bank, Development research Group, Working Paper 5430
- Knoben J., Oelremans L.A.G. (2006) “Proximity and Inter-Organizational Collaboration: A Literature Review” *International Journal of Management Review*, volume 8-2.
- Laursen K. & Reichstein T. & Salter A. (2011) "Exploring the Effect of Geographical Proximity and University Quality on University-Industry Collaboration in the United Kingdom" *Regional Studies*, Taylor & Francis Journals, vol. 45(4), pages 507-523.
- Pollitt, C. e Bouckaert, G. (2004) *Public Management Reform: A Comparative Analysis*, II edition, Oxford, Oxford University Press
- Rallet, A. and Torre, A. (1999) “Is geographical proximity necessary in the innovation networks in the era of global economy?” *Geojournal* 49, 373-380.



Beyond funding: What can acknowledgements reveal about credit distribution in science?¹

Adèle Paul-Hus^{*}, Adrián A. Díaz-Faes^{**}, Nadine Desrochers^{*}, Rodrigo Costas^{***}, Maxime Sainte-Marie^{*}, Benoît Macaluso^{****} and Vincent Larivière^{****}

^{*} adele.paul-hus@umontreal.ca, nadine.desrochers@umontreal.ca, msaintemarie@gmail.com;

École de bibliothéconomie et des sciences de l'information, Université de Montréal
PO Box 6128, Downtown Station, Montreal, Quebec, H3C 3J7 (Canada)

^{**} adrian.arias@cchs.csic.es

IFS, Spanish National Research Council (CSIC), Albasanz 26-28, 28037 Madrid (Spain)

^{***} rcostas@cwts.leidenuniv.nl

Centre for Science and Technology Studies (CWTS), Leiden University, P.O. box 905, 2300 AX Leiden (the Netherlands)

^{****} macaluso.benoit@uqam.ca; vincent.lariviere@umontreal.ca

École de bibliothéconomie et des sciences de l'information, Université de Montréal
PO Box 6128, Downtown Station, Montreal, Quebec, H3C 3J7 (Canada) and Université du Québec à Montréal,
Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Observatoire des Sciences et
des Technologies (OST), CP 8888, Succ. Centre-Ville, H3C 3P8 Montréal, Qc. (Canada)

ABSTRACT

Funding acknowledgements found in scientific publications have been used for decades to study the role of funding in science production. However, beyond funding information, acknowledgements convey the indebtedness of authors to individuals, institutions and organizations that contributed, in some way, to the research that lead to publication. The objective of this paper is to explore the different types of contributions acknowledged in WoS funding acknowledgement (FA) texts. The Correspondence Analysis performed in this study reveals that FAs offer a unique window on research and collaborative practices, credit distribution, and how these vary across disciplines. FAs thus contribute to make the traditionally “invisible contributions” visible to the scientific community. Results presented in this study go further in demonstrating that acknowledgements are not confined to credit attribution, as they include disclosures of conflict of interest.

INTRODUCTION

Funding acknowledgements found in scientific publications have been used for decades to study the role of funding in science production (e.g., Crawford and Biderman, 1970; Harten and Hooten, 1992). However, beyond funding information, acknowledgements convey the indebtedness of authors to individuals, institutions and organizations that contributed, in some way, to the research that lead to publication. Acknowledgements reveal the hidden infrastructure that supports scientific research, showing how colleagues, tools, materials, and grants are mobilized in the context of scientific endeavour (Cronin, 2005).

¹ This work was supported by the Social Sciences and Humanities Research Council of Canada (SSHRC) and the Fonds de recherche du Québec – Société et culture (FRQSC), as well as by the South African DST-NRF Centre of Excellence in Scientometrics and STI Policy (SciSTIP).

Acknowledgements can therefore be perceived as credit for contributions and could be used to better understand collaboration and the division of labor in the scientific field.

The objective of this study is to explore the different types of contributions acknowledged in WoS funding acknowledgement (FA) texts. More specifically, this study aims to answer the following research questions:

- What types of contributions are acknowledged in FA texts?
- How do these types of contributions vary by discipline?

DATA AND METHODS

Data were retrieved from Web of Science (WoS) Science Citation Index Expanded (SCI-E), which includes FA (i.e., funding information but also the full acknowledgement text provided by authors for all other types of contributions). The FA corpus used in this study was generated by collecting all 2014 articles and reviews from Biology, Biomedical Research, Chemistry, Clinical Medicine, Earth and Space, Engineering and Technology, Mathematics, and Physics totaling 880,809 FA texts for as many papers. Discipline assignment was done using the National Science Foundation field and subfield classification of journals.

A term extraction procedureⁱ was performed on FA texts to extract nouns and noun phrases. Since acknowledgements are collected and indexed by WoS only if they include funding source information (Paul-Hus, Desrochers and Costas, under review), the presence of funding information was considered a common denominator. Funding-related terms were thus removed from the acknowledgement text in order to focus the analysis on other types of acknowledgements. Additionally, proper nouns were removed from the corpus using python packages *nltk* (Bird, Klein and Loper, 2009) and *pattern* (De Smedt and Daelemans, 2012).

A frequency score was generated for each extracted lexical items (nouns and noun phrases), providing the number of times the item appears in the corpus. A threshold of a minimum of five occurrences was applied for a remaining corpus of 5,770 distinct lexical items. Finally, a Correspondence Analysis (CA) was applied on the corpus following the procedure described in Díaz-Faes and Bordons (2014) and using a MATLAB package (Vicente-Villardón, 2014).

RESULTS

Table 1 shows that FAs are not evenly distributed across disciplines, both in absolute terms and in relation to the proportion of FA-bearing papers per discipline. The FA ratio varies between 54% (Clinical Medicine) and 82% (Biomedical Research) for an overall ratio of 69%.

Table 1. Presence of FA by discipline for 2014 papers

Discipline	All papers	FAs	% of paper with FA
Clinical Medicine	373,185	202,464	54%
Engineering and Technology	222,263	149,157	67%
Biomedical Research	180,245	148,177	82%
Chemistry	144,020	114,747	80%
Physics	124,619	94,344	76%
Biology	98,893	72,171	73%
Earth and Space	85,305	66,055	77%
Mathematics	50,377	33,694	67%
Total	1,278,907	880,809	69%

Using CA, lexical items were grouped in five clusters using k-means clustering and cosine similarity. A threshold of 700 occurrences was set for visualisation purposes. Retaining five axes, over 85% of variance is explained and all disciplines can be interpreted on plane 1-3 (Figure 1) and plane 2-4 (Figure 2)ⁱⁱ.

Cluster 1, where authors mainly show their gratitude for the technical help and assistance received, is mostly found in Biomedical Research. Cluster 2 is formed from lexical items related to technical assistance, access to facilities and resources, and discussions associated to research work and projects. This pattern is found in Physics, Chemistry, and Engineering and Technology.

Cluster 3 reflects suggestions and comments but in contrast to Cluster 2, these contributions seem to be more related to manuscript improvement rather than to the research process itself. Furthermore, field work, a specific form of data collection associated with Biology as well as Earth and Space and involving uncontrolled environments, also characterize that cluster.

Cluster 4 gathers lexical items linked to manuscript and editorial assistance as well as data analysis support. Authorship and potential conflicts of interest, which are important concerns in clinical studies given the consequences of fraud and unethical behaviour in that field, also emerge as strong factors in this cluster, led by papers in Clinical Medicine. Cluster 5 appears as the most peripheral one. Mostly pertaining to Mathematics, this cluster is characterized by lexical items referring mainly to authorship and intellectual debts associated to manuscript preparation.

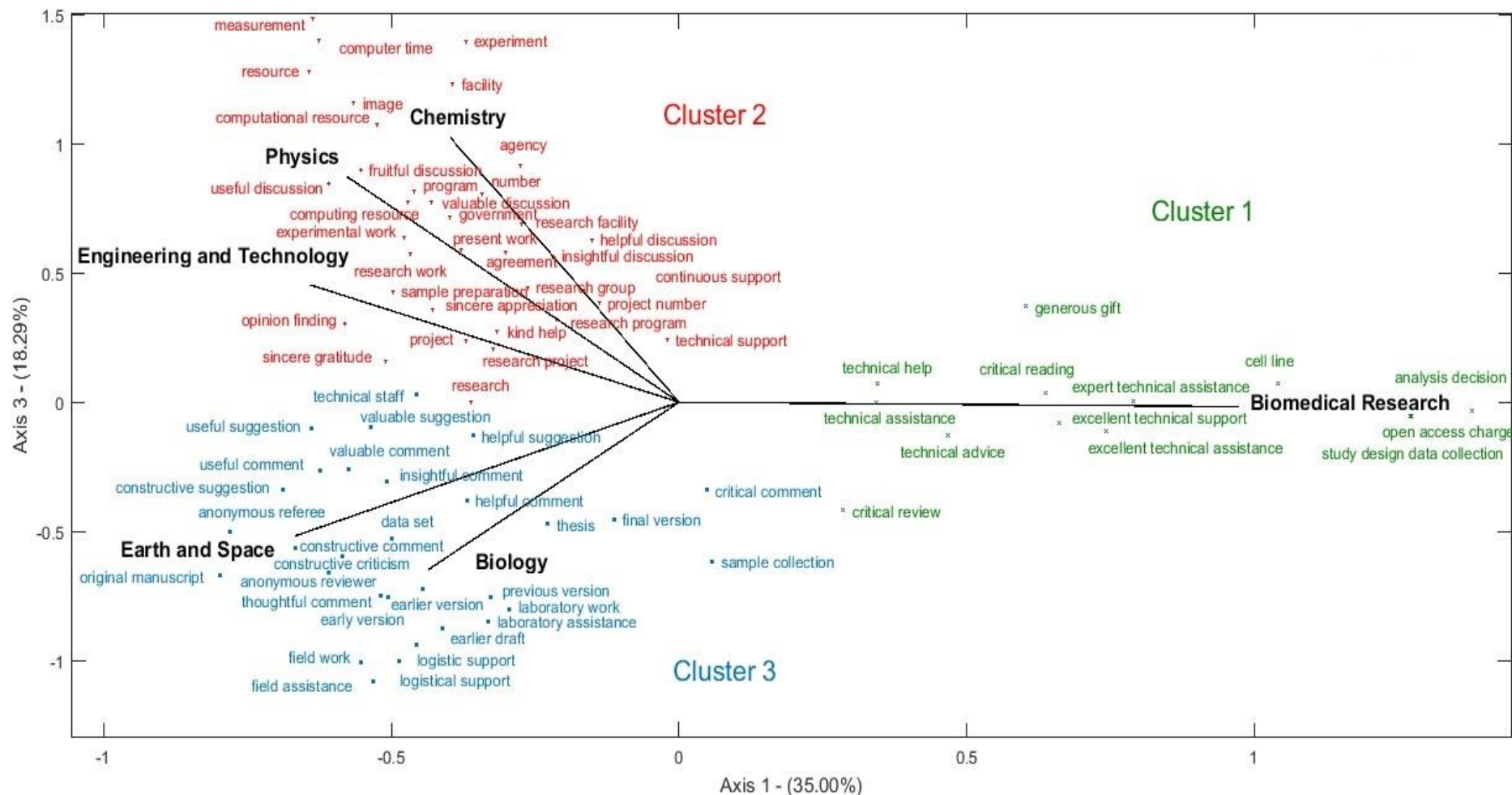


Figure 1. Bidimensional Correspondence Analysis for acknowledgements patterns by discipline (plane 1-3)

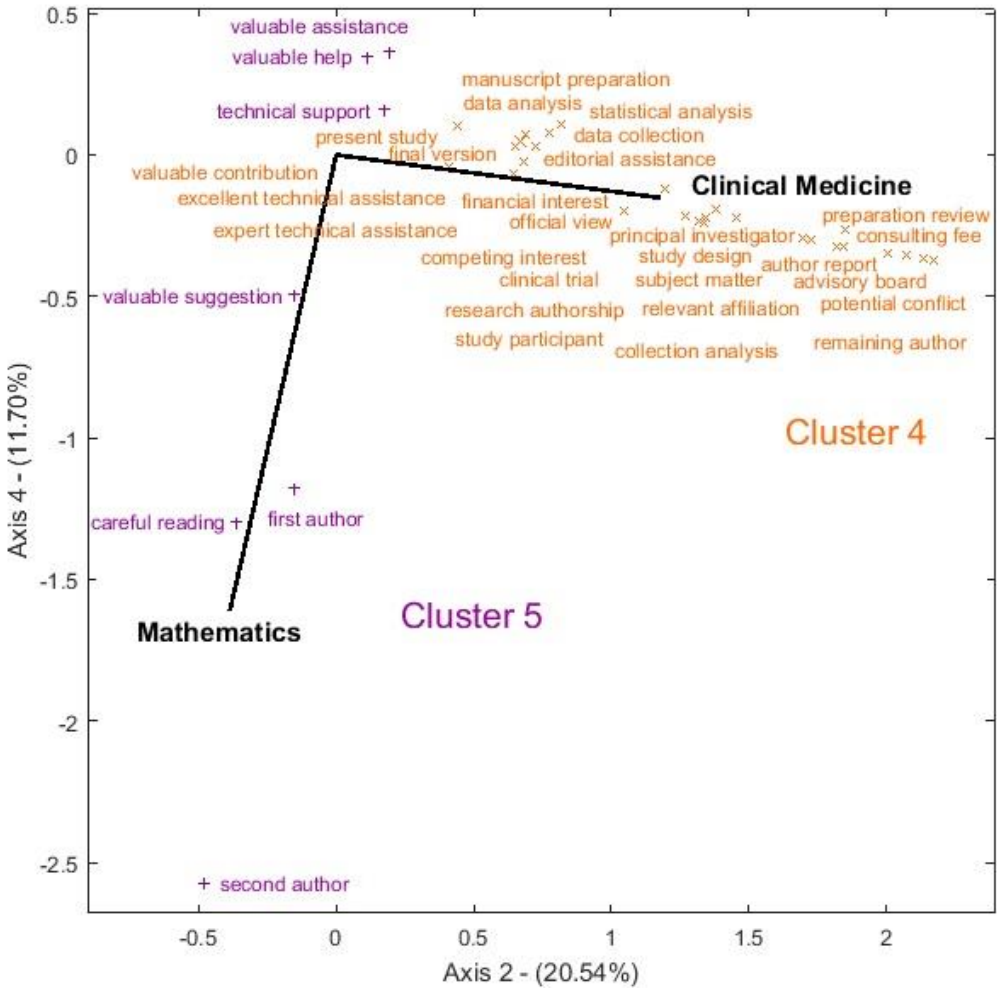


Figure 2. Bidimensional Correspondence Analysis for acknowledgements patterns by discipline (plane 2-4)



Table 2 presents the most frequent lexical items in the corpus. Percentages indicate the relative frequency of a given item for each discipline. Lexical items related to critical reading are among the most frequent, and two broad categories can be distinguished: acknowledgements to reviewers (e.g., “anonymous reviewer”) and those made to colleagues (e.g., “helpful discussion”), the latter being most frequently found in Physics. The remaining lexical items related to critical reading can be associated to either reviewers or colleagues (e.g., “helpful comment”).

As revealed by the CA (Figure 2), lexical items associated with conflicts of interest are mostly found in Clinical Medicine, in which disclosure of such potential conflicts is made mandatory by most journals’ guidelines (e.g., ICMJE, 2015). That being said, caution is required when analyzing terms out of context. For example, two frequent lexical items, “study design data collection” and “analysis decision”, do not constitute contribution acknowledgements *per se*, since most of their occurrences come from a conflict of interest disclosure statement found in journal guidelines, such as *PLOS* journals and *PeerJ*. The extracted items must thus be interpreted in the context of the original statement, such as, “The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.”



Table 2. Most frequent acknowledgement terms by discipline

Table 2. Most frequent FA lexical items by discipline

Acknowledgement terms	Biology	Biomedical Research	Chemistry	Clinical Medicine	Earth and Space	Engineering & Technology	Mathematics	Physics	Total
anonymous reviewer	12.8%	1.7%	0.3%	0.4%	20.9%	2.5%	2.5%	0.8%	3.7%
study design data collection	0.4%	19.2%	0.0%	1.2%	0.0%	0.0%	1.6%	0.0%	3.6%
analysis decision	0.4%	19.0%	0.0%	1.1%	0.0%	0.0%	1.6%	0.0%	3.6%
helpful discussion	1.1%	2.6%	2.7%	0.8%	2.8%	1.2%	1.5%	4.1%	2.0%
technical assistance	4.0%	3.4%	1.2%	2.2%	1.4%	0.7%	0.0%	0.7%	1.9%
helpful comment	4.2%	1.2%	0.3%	0.6%	5.2%	0.8%	3.9%	1.0%	1.5%
valuable comment	2.9%	0.6%	0.3%	0.2%	3.4%	1.2%	3.9%	0.8%	1.1%
technical support	1.8%	1.4%	1.0%	0.9%	1.6%	1.0%	0.0%	1.0%	1.1%
anonymous referee	2.2%	0.3%	0.1%	0.0%	5.4%	0.8%	6.4%	0.8%	1.1%
first author	2.0%	0.3%	0.1%	0.4%	1.8%	1.2%	9.4%	0.4%	1.1%
useful discussion	0.3%	0.5%	0.8%	0.1%	2.3%	0.7%	1.0%	4.6%	1.1%
constructive comment	2.3%	0.4%	0.1%	0.1%	6.5%	0.8%	1.7%	0.3%	1.0%
fruitful discussion	0.3%	0.4%	1.0%	0.1%	1.3%	0.7%	0.6%	2.8%	0.8%
research project	0.8%	0.5%	0.4%	0.4%	1.4%	1.0%	0.8%	0.6%	0.7%
critical reading	1.0%	1.9%	0.3%	0.6%	0.2%	0.1%	0.0%	0.4%	0.6%
data collection project	1.4%	0.7%	0.3%	1.2%	0.7%	0.1%	0.1%	0.1%	0.6%
useful comment	0.6%	0.5%	0.8%	0.1%	1.6%	0.5%	1.0%	0.4%	0.6%
official view	1.2%	0.3%	0.1%	0.1%	2.6%	0.2%	1.6%	0.9%	0.6%
earlier version	0.1%	0.8%	0.2%	1.4%	0.3%	0.1%	0.2%	0.1%	0.6%
valuable discussion	2.6%	0.3%	0.0%	0.2%	1.9%	0.2%	0.8%	0.1%	0.5%
second author	0.4%	0.4%	0.6%	0.1%	0.7%	0.5%	0.4%	1.5%	0.5%
research work	0.3%	0.0%	0.0%	0.0%	0.3%	0.5%	8.6%	0.2%	0.5%
analysis	0.4%	0.3%	0.6%	0.2%	0.6%	1.1%	0.4%	0.4%	0.5%
measurement	0.4%	0.4%	1.2%	0.1%	1.2%	0.5%	0.0%	0.1%	0.5%
measurement	0.0%	0.1%	1.8%	0.0%	0.1%	0.8%	0.0%	0.6%	0.5%

Color based on cell value: from darkest blue (lowest value) to darkest red (highest value)



DISCUSSION AND CONCLUSION

The CA performed in this study reveals that FAs offer a unique window on research and collaborative practices, credit distribution, and how these vary across disciplines. FAs thus contribute to make the traditionally “invisible contributions” visible to the scientific community, but the results here presented go further in demonstrating that acknowledgements are not confined to credit attribution, as they include disclosures of conflict of interest—or of their absence. These disclosures reveal that acknowledgements can also be self-declarations of ethical behaviour and tools to release third parties for any responsibility on the published results. Acknowledgements’ role might then be akin to that of contributorship statements and warrant another look in terms of their place in the evaluation of science and scientists.

Further steps in this study will include the addition of more disciplines (Social Sciences) and the use of linguistic processing techniques adapted to the idiosyncrasies of the FA corpus. This will allow for a better understanding of how acknowledgements can support the analysis of scientific practices beyond the core concern with funding.

REFERENCES

- Bird, S., Klein, E. and Loper, E. (2009). *Natural Language Processing with Python: Analyzing text with the Natural Language Toolkit*. Beijing: O'Reilly Media.
- Crawford, E. T., & Biderman, A. D. (1970). Paper money: Trends of research sponsorship in American sociology journals. *Social Science Information*, 9(1), 50–77.
- Cronin, B. (2005). *The hand of science: Academic writing and its rewards*. Lanham, Maryland: Scarecrow Press.
- De Smedt, T., & Daelemans, W. (2012). Pattern for python. *The Journal of Machine Learning Research*, 13(1), 2063-2067.
- Díaz-Faes, A. A., & Bordons, M. (2014). Acknowledgments in scientific publications: Presence in Spanish science and text patterns across disciplines. *Journal of the Association for Information Science and Technology*, 65(9), 1834-1849.
- Harter, S. P., & Hooten, P. A. (1992). Information science and scientists: JASIS, 1972–1990. *Journal of the American Society for Information Science*, 43(9), 583–593.
- ICMJE (2015). Uniform requirements for manuscripts submitted to biomedical journals: Ethical considerations in the conduct and reporting of research: Authorship and contributorship. <http://www.icmje.org/icmje-recommendations.pdf>
- Paul-Hus, A., Desrochers, N., and Costas, R. (under review). Characterization and description of funding acknowledgement data in Web of Science: considerations for their use. *Scientometrics*.

Vicente-Villardón, J.L. (2014). *MultBiplot: A package for Multivariate Analysis using Biplots*. Departamento de Estadística, Universidad de Salamanca [Software]. <http://biplot.usal.es/ClassicalBiplot/index.html>

ⁱ The term extraction transformation carried out by the SQL Server 2016 software : <https://msdn.microsoft.com/en-us/library/ms141809>

ⁱⁱ Lexical items and disciplines with variance explained below 40% are not displayed.



Allocating organisational level funding on the basis of Research Performance Based assessments, a comparative analysis of the EU Member States in international perspective¹

Koen Jonkers^{*}, Thomas Zacharewicz^{*}, Benedetto Lepori^{**}, Emanuela Reale^{***}

^{*} European Commission, Joint Research Centre, Institute for Prospective Technological Studies (IPTS), C/ Inca Garcilaso No. 3, E-41092 Seville, Spain, corresponding author: koen.jonkers@ec.europa.eu

^{**} Faculty of Economics, Università della Svizzera italiana, via Lambertenghi 10a, 6904 Lugano, Switzerland.

^{***} Research Institute on sustainable Economic Growth IRCRES, National Research Council CNR, Via dei Taurini 19, Rome, 00185 (Italy)

ABSTRACT

The paper analyses the extent to which RPBF allocation mechanisms are being implemented in Europe. To do so, this study builds on a novel set of data on project and organisational level funding developed for the European Commission, which identifies funding allocation mechanisms in each of the EU-28 Member States. This approach allows to compare the scope of RPBF systems across European countries. Further, the paper build on an in-depth analysis of RPBF implementation in 28 European countries, which comes to a classification of different types of RPBF implementation around three characteristics, i.e. a) the way research performance is measured and b) the type of link between performance assessment and allocation of resources. The analysis furthermore identifies a number of good practices while highlighting the potential for adverse effects of RPBF systems in research systems at different stages of development.

INTRODUCTION

In the EC communication on "Supporting Growth and Jobs – An Agenda for the Modernisation of Europe's Higher Education Systems" the European Commission recommends the introduction of *funding mechanisms linked to performance* which *introduce an element of competition* (EC, 2011). The Council Conclusions of November 2011 also promotes mechanisms linked to performance and competition to improve the governance of higher education systems.

Research performance based funding (RPBF) is defined as the allocation of organisational level funding to research organizations based on the (ex-post) assessment of their research performance (Hicks 2012). It is considered as one of the central tools through which many EU MS have tried to increase the effectiveness and performance of their Public Sector Research systems over the past decade.

The present paper aims to analyse the extent to which RPBF allocation mechanisms are being implemented in Europe, identifying strengths and drawbacks of different approaches. To do so, this study builds on a novel set of data on project and organisational level funding developed for the European Commission, which identifies funding allocation mechanisms in each of the EU-28 Member States. This approach allows to compare the scope of RPBF systems across European countries.

¹ This work was supported by the European Commission, DG JRC IPTS

Further, the paper build on an in-depth analysis of RPBF implementation in 28 European countries, which comes to a classification of different types of RPBF implementation around three characteristics, i.e. a) the way research performance is measured and b) the type of link between performance assessment and allocation of resources. In a future iteration of the paper also the amount of money affected by RPBF will be integrated. .

The analysis furthermore identifies a number of good practices while highlighting the potential for adverse effects of RPBF systems in, in the words of the conference theme, "central and more peripheral" European research systems.

Finally we aim to address what the main considerations, benefits and adverse effects of different types of research performance based funding systems are on national research systems.

Defining Research Performance Based Systems

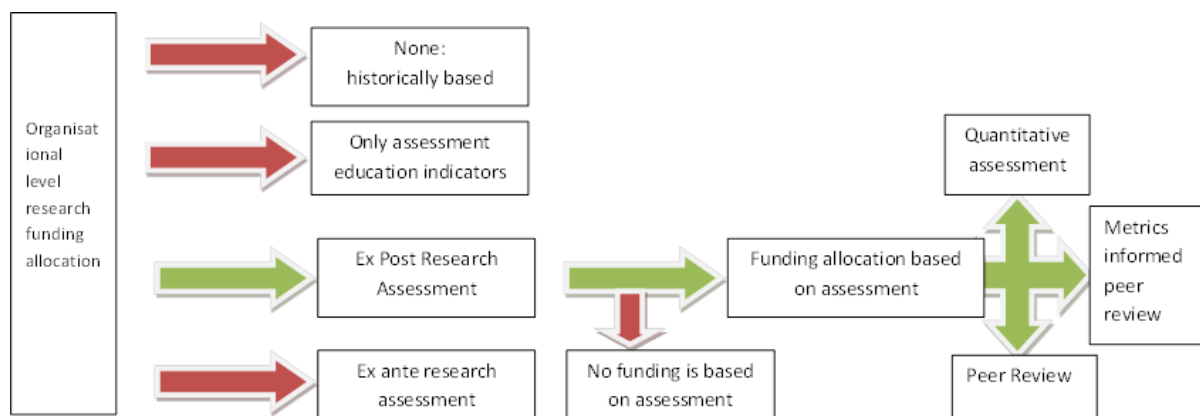
Public research funding is generally allocated in two main ways, through project funding and through organisational level funding (Lepori et al., 2007; van Steen, 2012). Considering the definition provided, "institutional funding" it may be more appropriate to talk about organisational level funding (Edquist, 1997; North, 1990). This definition has been operationalized in a series of statistical projects and data are now routinely collected by EUROSTAT at national level for a number of countries.

Organisational level funding for R&D can be allocated in different ways, based on historical considerations or negotiation between the State and the concerned institution or in a competitive manner (OECD, 2010). Building on Hicks (2012), research performance based funding systems are considered to be systems which base the allocation of organisational level funding for research (RPBF) on the basis of ex post assessments of research outputs.

This definition therefore excludes instruments which solely base their organisational level funding decisions on ex ante assessments such as the Excellenz Initiativ in Germany.

Figure 1 offers a schematic representation of organisational research funding allocation systems and delimitates the scope of RPBF in respect to other ways of allocating organisational level funding.

Figure 1 Schematic representation of institutional level funding allocation systems



Over the past decade many EU Member States have implemented RPBF systems, though the types of assessments and the share of resources allocated in this way differ widely. Many

countries use a funding formula which is partially based on the quantitative assessment of research outputs. Another set of countries rely instead on evaluations of research output through peer review. A subset of the latter allows these peer reviews to be informed by quantitative assessments of research outputs.

The following questions will be addressed in the paper:

- How is *ex-post performance assessed*? The existing literature suggests focusing in this respect on the distinction between metric-based systems and systems based on peer review (possibly informed by quantitative indicators)?
- What is the *nature of the link* between performance assessment and allocation of resources? A major distinction, in this respect would be between an automatic relationships (through a formula) and a discretionary relationships (for example through performance contracts)?
- What is the amount of resources allocated through competitive organisational level funding of which RPBF is a major subset?

METHODOLOGY AND DATA SOURCES

The approach to analyse the nature of organisational level funding allocation systems is twofold.

First, the scope of RPBF will be delimited through quantitative data collected in the context of the DG JRC funded PREF study. This study has engaged in the systematic collection of statistical data from national budgets, other administrative data and the accounts of research funders to assess the relative share of project funding and organisational level funding. Importantly, PREF provides a fine-grained division of organisational level funding, which is important since in most countries it is composed by streams with different characteristics. This will allow distinguishing RPBF-funding streams from other organisational level funding streams, where allocation is historical or based on the volume of educational activities, and measuring the amount of funding involved by RPBF.

Second, the paper will focus on a more qualitative assessment of the modalities of Performance Based Funding in the Member States. Through the information provided by a network of experts in the EU Member States, associated countries and selected third countries, 35 national research funding allocation mechanisms were examined according to a set of variables used to assess to which extent these countries implement RPBF systems as well as the nature of the underlying assessment.

These variables include: education metrics, the use of historically based funding allocation, bibliometric indicators (distinguishing between publication counts, journal impact based or citation based assessments) as well as other formula elements. Apart from the number of PhDs awarded, other (inputs or outputs) indicators frequently used refer to patent indicators, the participation in national or international research projects, external funding generated by contract research for companies or public administrations, income from Knowledge Transfer activities and spin off companies, gender composition of staff, and internationalisation indicators..

The paper then considers a number of bibliometric research output and impact measures to provide some information on the level of output of national research systems which could then be compared with the extent and characteristics of their RPBF.

RESULTS

The scope of organisational level funding

Data on the the share of project vs organisational level funding (table 1) display systematic differences in this respect between European countries, which allow distinguishing between three groups of countries:

- Countries where organisational level funding is dominant, like Italy and Spain.
- Countries where organisational level funding is more important, but project funding account for a significant share of public allocation, like Netherlands and Switzerland.
- Countries where the share of organisational level and project funding are similar, like the UK, Belgium and some Eastern European countries.

Table 1 share of organisational level and project funding in the EU28 MS

Country	year	Organisational	Project
AT	2013	72.7%	27.3%
BE	2014	47.6%	52.4%
BG	2014	56.4%	43.6%
CY	2014	79.0%	21.0%
CZ	2014	47.1%	52.9%
CH	2014	72.3%	27.7%
DE	2014	64.1%	35.9%
DK	2014	74.6%	25.4%
EE	2014	57.6%	42.4%
EL	2014	55.6%	44.4%
ES	2013	78.0%	22.0%
FI	2014	55.8%	44.2%
FR	2014	83.4%	16.6%
HR	2014	66.3%	33.7%
HU	2014	65.6%	34.4%
IE	2014	31.5%	68.5%
IT	2014	79.3%	20.7%
LT	2013	83.8%	16.2%
LU	2014	71.4%	28.6%
MT	2014	98.9%	1.1%
NL	2014	69.2%	30.8%
NO	2014	58,5%	41,5%
PL	2014	39.6%	60.4%
PT	2014	57.3%	42.7%
RO	2014	63.8%	36.2%
SE	2014	66.0%	34.0%
SI	2014	64.4%	35.6%
SK	2014	83.5%	16.5%
UK	2013	47.5%	52.5%

Note: ES: Data on regional funding not included; LV: not available

Source: preliminary (not yet validated) data from the PREF Dataset

On the one hand, it might be argued that the relevance of RPBF is higher in systems dominated by organisational level funding, where the competitive component of project funding is lacking; on the other hand, it should be investigated whether the RPBF and project funding are complementary or mutually supporting, i.e. those countries having high share of project funding also introduced RPBF to a larger extent. In some systems assessments of the success of research organisations in attaining project funding is used as one of the variables in research funding allocation formulae, thus reinforcing the effect on project funding decisions on the allocation of resources between organisations.

Performance-orientation of organisational level funding

Table 2 provides an overview of the different groupings of countries on the basis of the nature of the RPBF system they have in place.

Table 2. Characteristics of European RPBF systems

Country	No RPBF											Limited RPBF			Quantitative formula with Bibliometric assessment										Peer review				
	BG	CY	EL	ES	HU	IE	LU	LV	MT	RO	SI	AT	GE	NL	BE (FI)	BE (WA)	CZ	DK	EE	FI	HR	PL	SE	SK	FR	IT	LT	PT	UK
Education metrics	X			X	X	X	X			X		X	X	X	X	X		X	X	X				X					X
Historical	X	X	X	X	X			X	X	X	X						X	X										X	
Bibliometrics	Publications	X			X										X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Journal Impact Based	X					X								X		X	X	X	X		X					X	X	
	Citation	X				X									X	X					X		X	X	X	X			
Other formula elements	PhD graduates				X	X	X					X		X	X		X	X	X	X	X	X	X		X		X		
	Patents						X										X	X	X			X	X		X	X	X		
	Project funding					X							X			X	X	X	X	X	X		X	X	X	X			
	Business funding					X	X												X	X	X	X	X			X	X		
	Gender/diversity												X		X	X													
	Internationalisation												X							X	X			X	X	X	X		
Peer review												X													X	X	X	X	X
Performance Contracts												X		X				X	X	X									

The table categorizes the national research funding systems into four groups according to the criteria used for their allocation². The first one is composed of countries with no Research Performance Based Fundings, generally based on education metrics and historical considerations. The second group consists of countries with limited RPBF systems: i.e. without a clear ex post assessment of the (quality of) research output. This includes also systems which only take into account the granting of PhD degrees such as the Netherlands and several German Lander are considered to have a "limited RPBF system" since the production of doctoral theses is considered a research output – though the degree of "assessment" is limited. The countries classified as having an RPBF strongly vary in the mode of assessment of research outputs they employ.

The third category is composed of countries relying on quantitative formulas with bibliometric assessment to allocate research funding. The majority of countries base their

² Note that the number of indicators considered in the assessment exercises is not necessarily a sign of the level of sophistication of the system. For example, the peer review system in place in the UK is not deemed less sophisticated than the system in place in Italy.

funding allocation formulae on different types of metrics. Those which include a bibliometric assessment may adopt a journal based proxy for research impact such as Denmark, Norway and Poland, or adopt a citation based approach as in the case of Sweden or a combination thereof (e.g. Flanders). Among these bibliometric approaches there are considerable differences in the design of the assessment exercise. Some systems base themselves on counts of publications. Others weight these publications on the basis of the prestige or impact of the journals in which they are published either by using a journal impact factor measure (e.g. Poland) or by considering a differentiated list of journals selected and evaluated by expert panels (e.g. Denmark and Norway). Still another approach is to assess the impact of the publication output through citation analysis (e.g. Sweden) or a combination of the above approaches (e.g. Flanders).

Finally, the fourth group presents the countries mainly assessing research performance through peer review systems. The peer review systems implemented in the UK followed by Italy and Portugal are elaborate exercises in which a large number of organisational sub-units are assessed on the basis of submitted research outputs. There are however large differences in the extent to which the review outcome is linked to funding – the UK system generates from the review process a rating which has a very selective impact on funding and drives the allocation of the whole organisational level research funding to HEIs, while the other systems are less stringent and, like in the case of Italy, deal with only a limited portion of organisational level funding.

The latter remark emphasises the importance of focus not only on the existence of RPBF, but also on the nature and strength of the link between PRFB and funding allocation.

Effects of performance-orientation

Performance based funding, providing incentives for high impact output, is likely to have some effect on the level of excellence of the output of national science systems. The nature of the assessment on which funding allocation are based differs across countries (see also table 2). The approach and criteria adopted in these assessments provide different types of incentives. Apart from the potential positive effect of these incentives on e.g. the level of excellence of the national research output, there are known to be a number of potential adverse side effects. Peer review is often associated with potential conservatism, myopia, subjectivity and nepotism. Furthermore peer review systems can be relatively expensive (Geuna and Piolatto, 2015). Systems that rely on bibliometrics can e.g. incentivize gaming and sub-optimal publication behaviour. Quantitative publication and citation analyses also still suffer from a relative lack of acceptance from the academic community.

The available evidence on the effect of the different types of Performance Based Funding Systems is mixed. We find that most European systems have increased their performance on the bibliometric indicators considered (including e.g. the share of highly cited publications). Since most public research budgets have remained relatively stable or increased over the past decade (though a decrease is visible after 2008 in some countries) this is likely to be due to other factors. Potential explanations may include the growing Europeanisation or Globalisation of scientific fields (Nedeva and Wedlin, 2015). Institutional changes, including the introduction of RPBF are also likely to have played a role. There are systems without a clear RPBF system which perform very well on the bibliometric indicators considered (e.g. the Netherlands and Switzerland). These systems tend to have followed institutionally rooted

historical growth trajectories of scientific development going back decades of sustained and stable funding and gradual co-development of science, higher education and industrial development. Another potential explanation is that these systems tend to have adopted an alternative way of concentrating resources in top performing organisations: e.g. the binary university systems in place in The Netherlands and Switzerland (Lepori & Kyvik, 2010). The dominant publishing language and the fact that, as small highly internationalised countries, their publications are made to a relatively large degree in collaboration with international partners also affects their relative performance on such measures (OECD 2015). All the EU Member States which did not experience a consistent improvement in impact scores over the decade studied, did not have a RPF system in place. These countries, including Bulgaria, Romania and Latvia also received recommendations by international organisations to introduce RPF systems in recent years. There may however be other explanations for this relative under performance including chronic underfunding and the mobility of many of their best scientists to Western Europe and the US.

DISCUSSION AND CONCLUSION

This paper presents for the first time a comprehensive overview of the relative share of project and organisational level public funding for research in all EU28 Member States. In addition (to be included in a future iteration) the paper presents data on the share of competitively allocated organisational level funding from the PEF study. The latter is an indicator of particular policy relevance as EU policy makers increasingly recognize the multi-faceted nature of competitive funding and the extent to which it can be used as a policy lever for increasing the performance of national research systems.

The paper explores the different ways in which European Member States have implemented performance based funding regimes. The European research systems can be grouped in four categories according to the type of performance based funding they have implemented: a group of countries without RPF, a group of countries with limited RPF, a group of countries in which the RPF systems uses formulae based on quantitative indicators and a group of countries in which the RPF system uses formulae based on peer review. The latter may be informed, or not, by quantitative indicators.

The follow-up analysis will be developed in two directions: a) providing a more precise delimitation of RPF – related funding in terms of their specific streams and the amount of money involved and b) developing a more systematic characterisation of RPF along the three dimensions of the way performance is measured, the type of linkage between performance assessment and funding and the strength of this linkage.

Some of the best performing countries a different approach to the concentration of resources. Many other countries have introduced RPF relatively recently and it is therefore too early to assess their impact. Systems which have received recommendations in recent years to introduce performance based funding systems by the European Commission and other international organisations but have not yet done so, are the only countries which do not show a consistent improvement on the indicators considered.

Factors which are likely to influence the relative effect, acceptance and success of Performance Based Funding regimes include the share of organisational level funding which is allocated through RPF, the speed within which the system is introduced, the degree of

stakeholder involvement, the impact different systems have on the autonomy of research performers, the criteria on which they evaluate and their likely impact on research excellence indicators as well as the other missions and behaviours which the government wants to promote in these organisations.

REFERENCES

Council of the European Union (2011), Council conclusion on the modernisation of higher education, retrieved 14/03/2016 from:

http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/educ/126375.pdf

Cruz-Castro, L., Bleda, M., Derrick, G.E., Jonkers, K., Martinez, C., Sanz-Menendez, L., 2011. Issue brief.: Public sector research funding. OECD Innovation Policy Platform, retrieved 14/03/2016 from

https://www.innovationpolicyplatform.org/sites/default/files/Public%20Sector%20Research%20Funding_0_0.pdf

Edquist, C., Johnson, B., Institutions and Organisations in Systems of Innovation, in C. Edquist, *Systems of Innovation, Technologies, institutions and organisations*, (pp 41-60), Cheltenham: Edward Elgar

European Commission (2011), Supporting growth and jobs an agenda for the modernisation of Europe's higher education systems from:

http://ec.europa.eu/education/library/policy/modernisation_en.pdf

Geuna, A., Piolatto, M. (2016), Research assessment in the UK and Italy: Costly and difficult, but probably worth it (at least for a while), *Research Policy*, 45 (1), 260–271

Hicks, D., (2012). Performance-based university research funding systems, *Research Policy*, 41, (2), 251–261

Lepori, B., Benninghoff, M., Jongbloed, B., Salerno, C. & Slipersaeter, S., (2007). Changing models and patterns of higher education funding: Some empirical evidence. In A. Bonaccorsi & C. Daraio (Eds.) *Universities and Strategic Knowledge Creation. Specialization and Performance in Europe* (pp. 85-111). Bodmin, Cornwall: MPG Books Limited.

Lepori, B., Kyvik, S., (2010). The Research Mission of Universities of Applied Sciences and the Future Configuration of Higher Education Systems in Europe, *Higher Education Policy*, 23, 295–316.

North, D., (1990), *Institutions , institutional change and economic performance*, Cambridge (UK) Cambridge University Press

OECD, (2010). *Performance-Based Funding for Public Research in Tertiary Education Institutions*, Workshop Proceedings, Paris: OECD publishing

OECD, (2015), *Science, Technology and Industry Scoreboard*, Paris OECD Publishing

Steen, J. v., (2012). “Modes of Public Funding of Research and Development: Towards Internationally Comparable Indicators”, OECD Science, Technology and Industry Working Papers, 2012/04, Paris: OECD Publishing.
<http://dx.doi.org/10.1787/5k98ssns1gzs-en>.

CHAPTER 8

Collaborations, Mobility and Internationalisation



Predicting STEM Career Success by STI Knowledge Utilization Patterns

Barry Bozeman^{*}, Jan Youtie^{**} and Stuart Bretschneider^{***}

^{*} bbozeman@asu.edu

Arizona Centennial Professor of Technology Policy and Public Management
Center for Organization Research & Design (CORD)
Arizona State University | School of Public Affairs
411 N. Central Ave. Suite 480, Mail Code 3720, Phoenix, AZ 85004-0687, USA

^{**} jan.youtie@innovate.gatech.edu

Enterprise Innovation Institute
Georgia Institute of Technology | School of Public Policy
75 Fifth Street, N.W. Atlanta, GA 30308 USA

^{***} sbretsch@asu.edu

Foundation Professor of Organization Design and Public Management,
Center for Organization Research and Design (CORD),
Arizona State University | School of Public Affairs
411 N. Central Ave. Suite 480, Mail Code 3720, Phoenix, AZ 85004-0687, USA

ABSTRACT

As a part of discussion on knowledge utilization on science and technology, the mixed of papers presented in the panel discussion is designed to illustrate the patterns of collaboration, mobility, and diffusion of knowledge as well as those of labor force. In particular, the first two papers presented in the panel explore the potential of STEM career success through cosmopolitan collaboration and international community collaboration (focused on the relationships between China and Russia) in nanotechnology, which would provide implications on national and international benchmarking of innovation. For policy implications on graduate education and innovation, mobility pattern of non-U.S. Ph.D. degree holders is examined, and impact of a policy report on the target academic communities is investigated through development of credibility map. This panel is designed to highlight a recent effort of understanding geographical, cognitive or social spaces that are present in the scientific and technological activity as well as in doctoral education. The papers presented in this panel, therefore, will provide a rich set of significant and relevant insights drawn from examining STI knowledge utilization patterns to the STI-ENID community. The anticipated length of the event may be 90 minutes and there is no preferred number of attendees in particular although it is expected to be in between 35 and 60 at the minimum.

Paper 1: Career Impacts of Cosmopolitan Collaboration

Barry Bozeman, Arizona Centennial Professor of Technology Policy and Public Management, Center for Organization Research and Design, School of Public Affairs, Arizona State University, bbozeman@asu.edu

Monica Gaughan, Associate Professor, School of Human Evolution and Social Change, Arizona State University

Jiwon Jung, Senior Research Associate and PhD student at the Center for Organization Research and Design and School of Public Affairs, Arizona State University

Abstract

In previous studies, the concept “collaboration cosmopolitanism” has referred to the heterogeneity and social distance characteristics of academic researchers’ collaborators. Thus, collaborations with persons in one’s own department or laboratory would represent low levels of cosmopolitanism and collaborations with persons in different nations, in different disciplines or different sectors would represent higher levels of cosmopolitanism. Bozeman and Corley (2004) found that most researchers have low levels of CC but those having higher levels tend to be more successful in winning grants for their research. Lee and Bozeman (2005) found that CC predicts publishing productivity.

To this point, CC studies have focused on doctoral level researchers and in particular academic researchers. The current study seeks to extend the CC general hypotheses to a set of non-doctoral level research personnel, most of who work in industry. We first ask, “what career routes lead to cosmopolitan collaboration versus more parochial collaboration?” We then turn to capacity issues related to gender and race. We hypothesize that women and underrepresented minorities will tend to have less cosmopolitan collaboration patterns, due to diminished professionally relevant social capital and, in turn, that these diminished levels of CC will be negatively associated with positive career outcomes.

Our study employs National Science Foundation data, for which the present researchers are grant awardees and licensees; specifically, we use the 2006 and 2010 National Survey of College Graduates (NSCG) to measure CC, comparing doctoral and pre-doctoral workers. Focusing on changes between 2006 and 2010 (using NSCG data from both periods) the study examines impacts of CC on five distinct measures of career success: number employees supervised, job satisfaction, promotion, salary and skill augmentation.

Paper 2: Bounded Collaboration and Changing Core-Periphery Relationships in Sino-Russian Scientific Co-Production

Abdullah Gök, Manchester Institute of Innovation Research, Alliance Manchester Business School, University of Manchester, abdullah.gok@manchester.ac.uk

Maria Karaulova, Manchester Institute of Innovation Research, Alliance Manchester Business School, University of Manchester, maria.karaulova@postgrad.mbs.ac.uk

Philip Shapira, Manchester Institute of Innovation Research, Alliance Manchester Business School, University of Manchester and Georgia Institute of Technology,
PShapira@manchester.ac.uk

Abstract

Scientific collaboration between nations has always been considered one of the main science indicators. However, most of the attention is on the collaboration with or between centres of influence in science – countries that amass research outputs, publish main outlets for these outputs and set the general rules for academic conduct, including the language of publication. The dynamics of the collaboration between other countries and regional centres of influence has often been overlooked. In this paper, we investigate the bounded scientific co-production between countries in transition from periphery to core and vice versa. We also look into the influence of the relationship between core and transitional countries on the bounded collaboration between transitional countries, by employing a global systems perspective.

Empirically, we study the case of scientific collaboration between China and Russia in nanotechnology between 1990 and 2012. Over the past 100 years, the patterns of scientific co-production shifted significantly between the two countries. While China rose the most dynamically developing country in the world, the role of Russia declined from the core player to a peripheral actor in the regional and global research system. Previous research indicates that the rise of China's competence in nanotechnology has led to it becoming one of the global hubs of nanotechnology research and commercialisation (Shapira and Wang, 2010). Our research builds on these findings, reflecting China's gradual transition to become the leading country in terms of number of nanotechnology publications, ranking the first and producing about a quarter of the global output, while Russia has gradually declined from being one of the top actors to the 13th rank with around 3% share of the global output. Although there are considerable geographic, economic, cultural and historical proximities between China and Russia, their scientific co-production is significantly bounded. For Russia, China is the 19th largest international partner with a share of around 2% of all internationally collaborated publications. For China, Russia is also ranked as the 19th representing less than 1% of all international collaborations. For both countries leading international partners are the USA and major European countries. Our empirical focus is on the causes and dynamics of this phenomenon by utilising a range of indicators.

The conceptual framework of this paper derives from the world-system theory in which the relationship between core, periphery and semi-periphery countries is explained in reference to the dynamics of a global system to which they are embedded and with a special emphasis on path-dependencies. Our theoretical contribution is twofold. First, by applying the world systems theory to the study of scientific publication in a dynamic setting, we contribute to this theory which is often criticised to be neglecting socio-cultural production by over-emphasising economic production. Furthermore, we contribute to a global extension into the systems of innovation idea which is mostly considered in national and regional contexts.

Paper 3: Going home: why do non-US citizens with US Ph.D. degrees return home?

Stuart Bretschneider, Foundation Professor of Organizational Design and Public Management Center for Organization Research & Design, School of Public Affairs, Arizona State University

Abstract

Over the past 40 years the proportion of non-US citizens who receive Ph.D. degrees from US universities has grown from under 20% to almost half of all degrees awarded per year. During that same time period policy makers in the US and around the world have become more and more focused on how graduate education in general and doctoral education specifically fuels innovation and economic growth. As a result there has been growing concern about the career path of non-US citizens with US Ph.D. degrees and whether they are returning home. While many of these degree holders received both direct and indirect subsidies, a more salient concern among policy makers has been how decisions to return home might have long run consequences to the relative competitiveness of the US in the world economy. Previous work focusing on the rate at which these individuals stay suggests less than 30% return home (Finn 2014) but there is huge variation by country of origin and field of study. This paper makes use of data on individuals holding US Ph.D. degrees living in the US (NSF Survey of Doctoral Recipients) and outside the US (NSF International Survey of Doctoral Recipients) to develop and test an explanation for individual decisions to return home. The model includes the effects of labor market factors, individual and family effects, and policy variables that look at how individuals funded their doctoral education as well as efforts by their home countries to enhance their scientific and engineering infrastructures. Results suggest that in at least one country, efforts to actively attract these students home have had some effects at increasing their likelihood of return.

Paper 4: The Credibility of Policy Reporting Across Learning Disciplines

Jan Youtie, Enterprise Innovation Institute, Georgia Institute of Technology,
jan.youtie@innovate.gatech.edu

Abstract

The notion of a credibility map argues that everyone has a distinctive map that dictates the preference given to different types and sources of information. When seeking to influence other academic fields, scholars will likely turn to scientific and technical information though other types, such as policy reports, may also be relevant. We draw on the credibility map concept to understand how a major policy report is taken up by the target academic community. The report, *How People Learn*, was published by the US National Academies in 2001, to expose the education community (mainly educational researchers but also knowledge-seeking practitioners) to major cognitive science research findings of relevance to learning. Cognitive science is not unrelated to education, which would likely facilitate this diffusion. *How People Learn* summarized decades of cognitive science findings about learning from hundreds of studies into eight categories that address learners and teachers. We applied several search strings to measure the take up of this report in the target community. We used Google Scholar to evince that that *How People Learn* was cited in nearly 15,000 publications, these citations grew particularly steeply from 2000 to 2008, and most were in education-related journal articles and other publications. We performed a similar analysis using the Web of Science. While, there were only 300 Web of Science journal articles that cited this report, again, by articles mostly in education research journals, most of the citations were substantial as opposed to perfunctory. We employ an analysis of the number of citations, position of citations in the middle or later parts of the article, and coding of the content around the citations to measure the credibility and diffusion of this report.

REFERENCES

Bozeman, Barry, and Elizabeth Corley. "Scientists' collaboration strategies: implications for scientific and technical human capital." *Research Policy* 33, no. 4 (2004): 599-616.

Finn, M. G. (2014). Stay rates of foreign doctorate recipients from US universities, 2011, Oak Ridge Institute for Science and Education.

Lee, Sooho, and Barry Bozeman. "The impact of research collaboration on scientific productivity." *Social studies of science* 35, no. 5 (2005): 673-702.



The Credibility of Policy Reporting Across Learning Disciplines

Stephen Carley*, Jan Youtie**, Gregg Solomon***, Alan Porter*

* School of Public Policy, Georgia Institute of Technology, Atlanta, GA 30332 USA

**Enterprise Innovation Institute, Georgia Institute of Technology, Atlanta, GA 30308 USA

*** Directorate for Education and Human Resource, National Science Foundation, Arlington, VA 22230 USA

Correspondence: jan.youtie@innovate.gatech.edu

ABSTRACT

The notion of a credibility map argues that everyone has a distinctive map that dictates the preference given to different types and sources of information. When seeking to influence other academic fields, scholars will likely turn to scientific and technical information though other types, such as policy reports, may also be relevant. We draw on the credibility mapping concept to understand how a major policy report is taken up by the target academic community. The report, *How People Learn*, was published by the US National Academies in 2000, to expose the education community (mainly educational researchers but also knowledge-seeking practitioners) to major cognitive science research findings of relevance to learning. We applied several search strings to measure the take up of this report in the target community. We used Google Scholar to evince that that *How People Learn* was cited in nearly 15,000 publications, these citations grew particularly steeply from 2000 to 2008, and most were in education-related journal papers. We performed a similar analysis using the Web of Science, which showed that most of the citations were substantial as opposed to perfunctory.

INTRODUCTION

How People Learn (HPL) is a US National Academies report published in 2000 to improve US education by acquainting the education community with findings in the field of cognitive science the most effective ways to teach and learn. The idea behind this report was that the education researchers and practitioners were unfamiliar with the findings in cognitive science that relate to education, and that this gap was contributing to a decline in the ability of educators to teach and students to learn.

The extent to which this work has been taken up by its target audience is a matter of question. Of special interest to the present inquiry is whether HPL is deemed credible to the target audience (i.e., the education community) as well as other audiences. The question that follows is: how can we best assess if either audience is being influenced by this work?

This analysis measures the credibility of the report based on citations. It is acknowledged, however, that other techniques for measuring credibility exist. A case can be made that influencing a large percentage of a target audience (which happens to be numerically small in absolute terms) and a small percentage of a general audience are preferable and therefore, a measure of the credibility of the report. The results will show that the report is taken-up by the target education audience and that references to the report are not perfunctory.

LITERATURE REVIEW

Citations-based indicators enjoy a place of prominence in the evaluation of scholarly research today. Critics point to various faults with this indicator (Gingras and Wallace, 2010;

Carley et al., 2013; Carley and Porter, 2012). Despite its well-known imperfections, citations are the best indicator available in terms of measuring the take-up/credibility of literature. But while “citations are easily counted, it is not certain what is being measured” (Cano, 1989). It is also not certain that all citations are equal in credibility, with credibility referring to a combination of information and the judgements associated with this information (Bozeman, 1986). How is citation credibility best measured? Lutz Bornmann and Hans-Dieter Daniel provide a useful heuristic for categorizing citation types (Bornmann and Daniel, 2008). In Table 4 of their paper “What do citation counts measure? A review of studies on citing behavior” they outline 11 citation categories: (1) direct reference to an empirical finding in the cited document, (2) simple mention (of the type “compare here also,” “see also,” “see, for example”) without any further more specific reference to the cited document, (3) direct reference to a theory or concept in the cited document, (4) direct reference to a method in the cited document, (5) overview citation (of the type “for an overview, see here,” “see summary in”) without any further reference to the cited document, (6) use of a data collection method (such as a test) taken from the cited document, (7) word-for-word quotation of text in the cited document, (8) use of a statistical method taken from the cited document, (9) substantial, theoretical, or methodological critique of the cited document, (10) use of a table, figure, or list taken from the cited document, and (11) other citation type (for unclear citations). This taxonomy will be used to more precisely assess the citation credibility of HPL.

A further indicator of credibility is the location of the citation. In the article “Citation Behavior - Classification, Utility, and Location” Cano (1989) posits that citation location (within a citing article) provides utility as a bibliometric indicator, and one that has yet to receive a great deal of attention. Citation location can, in and of itself, provide guidance as to how substantive (or perfunctory) a given citation is likely to be. J.C. Smart builds on this classification by dividing articles into three parts: “one-third introduction and literature review, one-third research procedures and findings, and one-third discussion and implications” (Smart, 2005). This classification allows us to hypothesize and test the proposition that citations falling in the first third (or Introduction and/or Literature Review sections) of articles citing HPL are likely to be less substantive in nature than those falling in the final two thirds of these documents.

An additional indicator for citation impact is that of recurring citations—i.e. citations to the same work that occur more than once in a given citing article. The argument can be made that papers cited more than once are considered more credible (Hu, 2015). Citations receiving more attention are likely making more significant contributions to their citers than citations receiving a single and/or brief mention.

METHODOLOGY

The key question that this paper addresses, through an analysis of citation characteristics, is: how credible is HPL, especially to the education community? In essence, we are restating the challenge posed by Henry Small (2010) of the “conversion of non-citers to citers” (Small, 2010). The paper focuses on references to scientific literature as a measure of credibility because the formal study process that the US National Academies uses to develop its report, including HPL, places significant weight on review of the scientific literature, in conjunction with public meetings, third party submissions, and individual investigations (National Academies, 2006).

The primary databases from which citations to HPL are drawn and analyzed in this paper are the Web of Science (WOS) and Google Scholar (GS). These citation services, which continue to enjoy a significant amount of attention, provide an interesting study in contrast. Prior research has noted that WOS indexes more selectively (de Winter et al., 2014; Thomson Reuters, 2013a, as cited in de Winter et al., 2014) and is not accessible to the general public (de

Winter et al., 2014). GS, by contrast uses web crawlers to draw on a broader, less curated, range of materials (Google Scholar 2013). It has been argued that because of its automatic inclusion process, GS is susceptible to errors in metadata (Jacso' 2008) and to indexing of non-scientific works (Cathcart and Roberts 2005; Donlan and Cooke 2005; Jacso' 2005a; Vine 2006; Wleklinski 2005; de Winter et al., 2014).

WOS citations to HPL were collected using the WOS Cited Reference Search feature in April of 2015. GS citations to HPL were identified using the Publish or Perish¹ General Citation feature shortly thereafter. The query included "Bransford" in the author field and "How People Learn" as the search term. Results were restricted to those works published in 1999 or 2000. After removing false positives Zotero² is used to merge duplicates (duplicates are not merged if they come from different document types) and download GS citations to HPL in spreadsheet format. Results for all procedures were imported into VantagePoint³ and analyzed using this software.

RESULTS

Citation analyses indicate that HPL has been recognized by a fairly large audience. Canvassing GS for this report yields nearly 15,000 citations as of August 2015. GS cites are collected via Publish or Perish, with "Bransford" listed in the author field and "How People Learn" listed as the search term. Results are restricted to only those works published in 1999 or 2000. Fielded data include citing authors, titles, years, item types and publishers. While subject areas are not provided by GS citation results, keywords from citing titles can prove useful for getting a sense of the type of literature paying attention to HPL. We find that 14% of GS citations to HPL contain the term "education" in their title. We note from that GS citations to HPL follow an uneven growth trajectory: in the 1999-2009 period citation increase might be described as one of steady (and at times rapid) growth, while a leveling off period may be occurring from 2009 onward.

Although the report has broad take-up, what is its credibility relative to education researchers? As of April 2015 there were over 300 citations to HPL indexed in WOS. These span 16 publication years, nearly 90 Web of Science Categories, nearly 40 countries and 212 citation sources (most, but not all, of these are journals). The average WOS citation trend to HPL is one of increase (although growth patterns in this figure might be described as uneven as well). We note as well that WOS citations are not always immediately accounted for in real-time, but citation counts can be backfilled to reflect earlier activity. We also observe that citation counts indexed on GS dwarf those indexed on WOS – by a factor of 40. Is this result atypical? As has been previously mentioned, de Winter and colleagues (2014) indicate that social science publications are expected to attract more citations on GS than they are on WOS. These authors compare citation counts for a 1955 article by Eugene Garfield and find that GS citations grow at a significantly faster rate for this work than do WOS citations. In light of these findings the HPL citation differential between WOS and GS is not unexpected.

Among the more than 300 WOS citations to HPL, three-quarters fall into an education related WOS Category.⁴ Similar to the GS finding, just over 15% of WOS citations to HPL contain the word 'education' in their title. Given our interest in the credibility that HPL has with the education community, we test whether the citations to HPL are of a more perfunctory or substantive nature. To this end, we were able to obtain full text records of two-thirds of the WOS articles with cited references to HPL, which formed the basis of our credibility analysis.

¹ see www.harzing.com/pop.htm

² see www.zotero.org

³ see www.thevantagepoint.com

⁴ The four education related WOS Categories that cite *How People Learn* are: (i) Education & Educational Research, (ii) Education, Scientific Disciplines, (iii) Psychology, Educational, and (iv) Education, Special.

Sixteen percent of these citing articles contained recurring citations. The number of recurring references to HPL steadily increased over time. By this indicator HPL is steadily growing in influence. Moreover, among the more than 30 recurring references to HPL, roughly 80% are affiliated with an education based WOS Category⁵. By this metric HPL is steadily growing in influence with target audiences.

As has been previously mentioned, citation location can provide a telling indicator of the credibility of the report. Citations fall in the final two thirds of a paper are more likely to be of a substantive (as opposed to perfunctory) nature (Hu, 2015). To test this proposition, we first identify substantive citations by classifying them according to the taxonomy provided in Table 4 of Bornmann and Daniel (2008, p. 55). Using the scheme in their paper we can match citation type to citation location. The latter is accomplished by dividing the page a given citation occurs on by the total number of papers in the citing article. The citation types (i) ‘Overview citation (of the type “for an overview, see here,” “see summary in”) without any further reference to the cited document,’ and (ii) ‘Simple mention (of the type “compare here also,” “see also,” “see, for example”) without any further more specific reference to the cited document’ might generally be referred to as perfunctory citations, while the remaining citation types are more substantive in nature. Most of the substantive citations (i.e. 5 out of the 6 reference types) in this table have a mean location in the final two thirds of their citing documents. The mean location for all citations 0.38, indicating that the average location for all citations to HPL falls in the second third of citing articles – a location indicating substance and/or impact. We also note that the majority of citations by reference type belong to the group ‘Direct reference to a theory or concept in the cited document’. This suggests that the majority of citers to HPL are specifically interested in the theories and concepts it offers – a sign that this document is having a more substantive impact.

CONCLUSION

At the outset, it was suggested that the credibility of source of information could be investigated by examining citation to this source. Credibility would be strongest to the extent that the report reaches and is used, via citation, by a larger percentage of a target audience, but presumably a relatively smaller percentage of a general audience. The results of this study suggest that HPL has both reached its target audience, as represented by the WOS analysis, as well as a more general audience as represented by the GS analysis. The report appears to have credibility beyond the narrow policymaking body that developed it to the broader education user community. This conclusion is subject to several limitations. No interviews with the education community were conducted, which would have enhanced the degree of understanding of the credibility of that HPL has relative to other sources of information used by this community. Moreover, citations are not the same as use, as has been summarized in Bornmann and Daniel (2008). Although more work is necessary to understand the credibility of HPL, including extending this methodology to other National Academy reports as well as reports of other policy advisory agencies, this study does offer a methodology and quantitative results to weigh in on an area that has largely been assessed through anecdote and case study (Youtie et al., 2016).

ACKNOWLEDGEMENTS. This work was supported by the US National Science Foundation, Division of Research on Learning in Formal and Informal Settings (DRL-1348765) and Science of Science and Innovation Policy, (Award #1262251). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

⁵ The education based WOS Categories associated with recurring citations are: (i) Education & Educational Research, (ii) Education, Scientific Disciplines and (iii) Psychology, Educational.

REFERENCES

- Bornmann, L. and H.D. Daniel (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, 64 (1): 45-80.
- Bozeman, B. (1986). The credibility of policy analysis: Between method and use. *Policy Studies Journal*, 14(4), 519-539.
- Bransford, J.D. A.L. Brown, R.R. Cocking (2000). *How People Learn*, 2nd ed. Washington, DC: National Academy Press. Print.
- Cano, V. (1989). Citation behavior - classification, utility, and location. *Journal of the American Society for Information Science*, 40 (4): 284-290.
- Carley, Stephen and Alan L. Porter (2012). A forward diversity index. *Scientometrics*, 90 (2): 407-427.
- Carley, Stephen, Alan L. Porter and Jan Youtie (2013). Toward a More Precise Definition of Self-Citation. *Scientometrics*, 94 (2): 777-780.
- de Winter, Joost C.F., Amir A. Zadpoor and Dimitra Dodou (2014). The expansion of Google Scholar versus Web of Science: a longitudinal study. *Scientometrics*, 98 (2): 1547-1565.
- Garfield, Eugene (1955). Citation indexes for science: A new dimension in documentation through association of ideas. *Science*, 122 (3159), 108-111.
- Gingras, Yves and Matthew L. Wallace (2010). Why it has become more difficult to predict Nobel Prize winners: a bibliometric analysis of nominees and winners of the chemistry and physics prizes (1901-2007). *Scientometrics*, 82 (2): 401-412.
- Goodrum, Abby, Katherine McCain, Steve Lawrence, and C Lee Giles (2001). Scholarly publishing in the Internet age: a citation analysis of computer science literature. *Information Processing & Management*, 37 (5): 661-675.
- Hu, Zhigang, Chaomei Chen and Zeyuan Liu (2015). The Recurrence of Citations within a Scientific Article. Salah, A.A., Y. Tonta, A.A. Akdag Salah, C. Sugimoto, U. Al (Eds.), Proceedings of ISSI 2015 Istanbul: 15th International Society of Scientometrics and Informetrics Conference, Istanbul, Turkey, 29 June to 3 July, 2015, Bogaziçi University Printhouse., At Istanbul, Turkey.
- Kochen, Manfred. (1987). How well do we acknowledge intellectual debts? *Journal of Documentation*, 43: 54-64.
- National Academies (2006). Our Study Process: Ensuring Independent, Objective Advice. http://www.nationalacademies.org/site_assets/groups/nasite/documents/webpage/na_069618.pdf (accessed June 24, 2016).
- Publish or Perish (2015). Harzing. Retrieved May 1, 2015, from www.harzing.com/pop.htm.
- Small, Henry (2010). Referencing through history: how the analysis of landmark scholarly texts can inform citation theory. *Research Evaluation*, 19 (3), 185-193.
- Smart, J.C. (2005). Attributes of exemplary research manuscripts employing quantitative analyses. *Research in Higher Education*, 46 (4): 461-477.
- VantagePoint (2015). Search Technology. Retrieved September 1, 2015, from www.thevantagepoint.com.
- Youtie, Jan, Bozeman, Barry, Jabbehdari, Sahra, Kao, Andrew (2016). Credibility and Use of Scientific and Technical Information in Policy Making: An analysis of the information bases of the US National Research Council's committee reports. White Paper.
- Zotero (2015). Center for History and New Media (GMU). Retrieved August 1, 2015, from www.zotero.com.



Mobility in the academic careers at the Flemish universities – Results from the Human Resources in Research database

Noëmi Debacker, Karen Vandavelde

Research Department, Ghent University, Ghent, Belgium
Centre for R&D Monitoring (ECOOM) Ghent

ABSTRACT

Worldwide mobility of researchers is generally correlated with an enhancement of the continuous interchange of ideas and is therefore highly stimulated. Flemish universities have occasionally been criticized for providing limited career opportunities for foreign researchers. However, presently no reliable data is available on researchers' mobility in Flanders. For the five Flemish universities a full dataset is available at the ECOOM Centre for R&D Monitoring, containing all academic appointments and PhD's awarded since 1990-1991. This dataset allows for prospective analyses within the Flemish university system as well as retrospective analyses and provides information on international and interuniversity recruitment and academic inbreeding. The present results confirm that the academic system in Flanders is indeed recognized by a tradition of internal recruitment. However specific differences between the universities and the scientific disciplines are observed calling for a more complex study of researchers' mobility and of the relationship between mobility and research excellence.

BACKGROUND

For a number of decades, the promotion of researchers' mobility has been at the core of many governments' research policies, expected to create spill-over effects in terms of research quality and level of innovation (Seeber & Lepori, 2014). A considerable number of funding schemes as well as policy documents identify an open labour market for researchers as one of the pillars of an excellent science system (EC, 2014). Worldwide mobility of researchers is generally correlated with an enhancement of the continuous interchange of ideas and is therefore highly stimulated (Horta, 2013; Soler, 2001; Tavares, Cardoso, Carvalho, Sousa, & Santiago, 2015); (Horta, Veloso, & Grediaga, 2010). A European-wide study has attempted to map the intensity of researchers' mobility as well as mobility flows and hurdles (MORE2, 2010). Such studies are mostly qualitative, empirical, field-specific, or survey-based, and often limited in scope or susceptible to bias due to low response rates. Mapping the results of mobility initiatives on a more systemic level, or identifying the intensity of researchers' mobility in a particular region, is an exercise faced with many hurdles, in particular as mobile researchers are notoriously difficult to track. As such, lack of valid information and reliable data lead to numerous assumptions with respect to researchers' mobility.

Flanders is an interesting region to examine for a number of reasons. First, Flemish universities have occasionally been criticized for providing limited career opportunities for foreign researchers. The EUMIDA project, for example, identified relatively low levels of internationalisation when compared to universities in similar European regions (Lepori, Seeber, & Bonaccorsi, 2015). Second, systemic observation is feasible as for the five Flemish universities a dataset is available at the ECOOM Ghent, containing all academic appointments and PhD's awarded since the academic year 1990-1991. Third, in addition to being able to

measure the degree of internationalisation amongst staff, this database allows for retrospective analyses (e.g. the Flemish career history of professors currently appointed).

In this study we examine both the interuniversity (i.e. national) mobility and international recruitment (i.e. graduates who had a prior academic career outside Flanders) in Flanders.

METHOD

In 2007, ECOOM Ghent established a database Human Resources in Research Flanders (HRRF) designed to monitor PhD production and academic careers within the 5 Flemish universities. This database contains appointment data from 1990-91 onwards of all scientific staff affiliated to a Flemish university as well as PhD registration and completion data in Flanders. At present the database contains data up to the academic year 2012-13. The career track analyses in this paper are limited to researchers born after 1964; analyses are as such limited to the researchers who started their academic career from 1990-91 onwards.

In Flanders the promotion towards full professorship generally starts with the position of tenure-track professor, which is the logical career step following a post-doc for those establishing an academic career. Appointments from tenure-track professor onwards are considered stable academic career positions combining research and teaching under the umbrella name 'ZAP' (professorial staff). For the ease of the current paper we will continue using ZAP for the denotation of this group.

For the analyses in this paper we determined the 'dominant' university of the post-doctoral and the ZAP appointment, and the 'dominant' scientific discipline by using the duration and percentage of the various appointment lines.

A position was considered as ZAP when the total percentage of a ZAP appointment at a given time was 50% or higher. An exception was made for law and medicine, for which the cut-offs were 30% and 20% respectively, given that ZAP-staff in these disciplines often combine research and teaching with practice outside university and thus tend to have lower appointment percentages.

The universities in the results section will be denoted using their relative size: 2 large-size universities and 3 small to medium size universities (LSU and SMSU respectively).

Bivariate analysis is used for descriptive findings and multiple regression is used to control for confounding. Multinomial regression analysis is used when dealing with non-dichotomous outcome variables.

First we analyse the composition of the post-doctoral population affiliated to the Flemish universities with respect to their institutional origin defined as (1) PhD at the same university, (2) PhD at a different Flemish university and (3) PhD outside Flanders. Next we analyse the composition of the ZAP population with respect to their institutional origin defined as (1) entire career within the same university, (2) ≥ 1 switch within Flanders, (3) PhD outside Flanders, post-doc and ZAP in Flanders and (4) no academic history in Flanders.

RESULTS

We first look into the composition of the post-doctoral researchers affiliated to a Flemish university between 1990-91 and 2012-13.

Nearly 60% of post-doctoral researchers within Flanders obtained their PhD at the same university; 35% obtained their PhD outside Flanders; only 6% obtained their PhD at a different Flemish university. Of the group with a Flemish PhD, 92% was Belgian; in the group with the PhD outside Flanders only 18%; leading to a total of 66% Belgians among the post-doctoral researchers¹.

We observe significant variation between the universities ($\text{Chi}^2(8, n=7979)=455.73, p<0.001$) (Table 1). In one of the LSU's 71% of the post-doctoral researchers originated from the same university whereas this share is lower, and all within a similar range, in the other four universities (ranging from 51% to 55%). The share of post-doctoral researchers who obtained their PhD in a different Flemish university varies significantly between the universities (from 2.8% to 22%): the smaller universities recruit more researchers from other Flemish universities than the larger ones. Finally, with respect to the share of researchers who obtained their PhD outside Flanders, once again we observe large variation between the universities (from 24% to 43%): the larger universities recruit more from outside Flanders compared to the smaller ones. One exception however is the LSU highlighted earlier because of the large share of internal recruitment at the post-doctoral level.

With respect to scientific discipline we observe significant differences as well ($\text{Chi}^2(8, n=7950)=232.61, p<0.001$) (Table 1) - however, the differences as such are smaller. Post-doctoral researchers in the natural sciences are the most atypical with respect to their composition, mainly caused by a higher share of researchers who obtained their PhD outside Flanders (47%). In the remaining disciplines the share of holders of a non-Flemish PhD at post-doctoral level varies between 27% and 34%. The Flemish interuniversity mobility on the other hand is lowest in natural and applied sciences.

The group of female post-doctoral researchers contains more researchers who obtained their PhD at the same university (62.5% versus 56.8% for men) at the expense of researchers who obtained their PhD outside Flanders (31.1% versus 37.9% for men) ($\text{Chi}^2(2, n=7979)=38.39, p<0.001$).

In order to exclude whether the differences in population composition between the universities can be explained by their specific scientific profile or vice versa we use multinomial logistic regression in which gender, main university of post-doctoral appointment and scientific discipline are included as independent variables (Table 2). Irrespective of the scientific discipline and gender, the differences as observed in the bivariate analysis remain.

We now look into the composition of the ZAP-staff in Flanders. We identify four mobility patterns in the earlier career of ZAP-staff: (1) those affiliated with one single Flemish university for their entire career amounted to 63%; (2) 20% had their entire career within Flanders but with ≥ 1 university switch; (3) 12% did not have any previous academic appointment within Flanders; (4) finally, 6% of the ZAP-staff started their career in Flanders at post-doctoral level. The shares of Belgian researchers in these groups are respectively 97%, 96%, 25% and 56%, leading to a total of 86% Belgians among the ZAP-staff.

Broken down by the university of the ZAP-affiliation the share of researchers with an entire career track at the same university varies widely, from 39% to 72% ($\text{Chi}^2(12, n=1257)=122.99, p<0.001$) (Table 1). The larger universities have the highest shares of ZAP-staff who spent their entire career at the same university and they have a lower share of ZAP-staff with ≥ 1 university switch within Flanders. The recruitment of ZAP-staff with a prior academic career outside Flanders also shows some important variation: in one LSU the share is 22.8% compared to 11.6% in the other LSU.

¹ The HRRF database covers the universities in Flanders, the northern half of Belgium. The nationality of its inhabitants is however « Belgian », not « Flemish ».

Broken down by scientific discipline we again observe significant differences ($\chi^2(12, n=1255)=50.29, p<0.001$) (Table 1). Medical and applied sciences have the highest share of ZAP-staff who spent their entire career at the same university (73%). Especially social sciences have a large share of ZAP-staff with ≥ 1 university switch in Flanders. The highest share of ZAP-staff with a prior academic career outside Flanders is found in humanities, social and natural sciences.

There are no significant gender differences with respect to the institutional origin of ZAP (Table 1).

These findings remain in a multinomial logistic regression analysis with the university of the ZAP-affiliation and the scientific discipline as independent variables (Table 3).

DISCUSSION

These analyses confirm earlier findings that the academic system in Flanders is recognized by a strong tradition of internal recruitment and promotion. Among the post-doctoral researchers 59% obtained their PhD at the same university and 35% obtained their PhD outside Flanders. Among the ZAP-staff 63% obtained their PhD at the same university and only 17% started their career outside Flanders. Given that this study is limited to the population born from 1965 onwards we expect that the number of homegrown ZAP is even higher in the total population of ZAP-staff. Similarly to these results a recent Swedish report mentions that more than 50% of academics has received their PhD at the same institution (Stringfellow, 2016).

Important differences exist between the universities and between the scientific disciplines, each independently from one another. To some extent, the size of the university explains some of the differences: the three SMSU's more often tend to recruit researchers who have previously worked in another Flemish university. However university size seems less of a deciding factor when it comes to the recruitment of staff with academic origins outside Flanders. The results indicate the existence of fundamental differences between the universities. In some cases the degree of mobility is clearly related to **country-level characteristics** such as national regulations, geography, and the science system (Seeber & Lepori, 2014). The variation in international recruitment suggests that **institutional features** (e.g. recruitment policies, university prestige, structural components) also have an impact on mobility patterns. The extent to which **individual elements** such as personal aspirations and career 'sacrifices' play a role, is out of the scope of this study. The case of Flanders as such identifies the target of possible policy intervention: if researcher mobility in general requires stimulation (outgoing mobility, interuniversity mobility, international mobility), country-level characteristics need to be addressed. If only international recruitment is a particular policy aim, additional interventions will be most effective at institutional level.

The present data show that the HRRF database is unique in its provision to map the career track of the researchers working within the Flemish university system and to monitor changes over time. Recruitment indicators such as those produced by the HRRF can be easily generated as internationally comparative indicators. They are a valuable starting point to identify discrepancies between institutional/national strategies on the one hand and everyday hiring practices on the other.

Also interesting, however, is the observation that Flanders may be quite atypical in the sense that the internal recruitment system has not stood in the way of high performance, evident in research-based rankings, innovation indicators etc. (ECOOM, 2015). This may shed new light on the findings from earlier studies. Contrary to most other studies on research mobility, this observation suggests that internal recruitment is not necessarily correlated with poor performance. An explanation for this may lie in the fact that the Flemish university system is highly competitive, ensuring tough selection at the moment of recruitment, and continuing to

push researchers to their limits throughout their career. It would be interesting to examine whether in this context, policy interventions stimulating more researcher mobility still have a positive impact on the performance of the research system. For the moment, the data suggest there is no single benchmark indicating the « right » balance between homegrown academics and externally recruited ones.

REFERENCES

- European Commission. (2014). Researchers' Report 2014 - Final Report. Brussels: DG Research and Innovation.
- ECOOM. (2015). Vlaams indicatorenboek Wetenschap, Technologie en Innovatie. Leuven.
- Horta, H. (2013). Deepening our understanding of academic inbreeding effects on research information exchange and scientific output: new insights for academic based research. *Higher Education*, 65(4), 487-510.
- Horta, H., Veloso, F. M., & Grediaga, R. (2010). Navel Gazing: Academic Inbreeding and Scientific Productivity. *Management Science*, 56(3), 414-429.
- Lepori, B., Seeber, M., & Bonaccorsi, A. (2015). Competition for talent. Country and organizational-level effects in the internationalization of European higher education institutions. *Research Policy*, 44(3), 789-802.
- MORE2. (2010). Support for continued data collection and analysis concerning mobility patterns and career paths of researchers. Brussels: IDEA Consult.
- Seeber, M., & Lepori, B. (2014). The Internationalization of European higher education institutions. *Knowledge, diversity and performance in european higher education*, 138-166.
- Soler, M. (2001). How inbreeding affects productivity in Europe. *Nature*, 411(6834), 132-132.
- Stringfellow, A. (2016). Vetenskapsradet turns spotlight on Sweden's internal recruitment habit. *Research Europe*, 424, 17-17.
- Tavares, O., Cardoso, S., Carvalho, T., Sousa, S. B., & Santiago, R. (2015). Academic inbreeding in the Portuguese academia. [journal article]. *Higher Education*, 69(6), 991-1006.



Table 1. Composition of post-doctoral researchers and ZAP-staff with respect to their institutional origin broken down by university of affiliation, scientific discipline and gender – HRRF 1990-2012 – subjects born from 1965 onwards

	Pool of post-doctoral researchers			ZAP-staff					
	N	% PhD at same university	% PhD at other Flemish university	% PhD outside Flanders	N	% Entire career at same university	% At least one switch within Flanders	% PhD outside Flanders – post-doc and ZAP in Flanders	% No academic history in Flanders
University of affiliation									
LSU1	(1)	70.7	5.0	24.3	(1)	71.9	16.5	2.8	8.8
LSU2	(1)	54.7	2.8	42.6	(1)	68.2	8.9	7.7	15.1
SMSU1	(1)	54.8	8.1	37.1	(1)	57.6	26.0	6.8	9.6
SMSU2	(1)	51.8	11.1	37.2	(1)	43.1	40.3	6.6	9.9
SMSU3	(1)	51.2	22.2	26.6	(1)	38.5	35.4	9.2	16.9
Total	7979	58.9	5.7	35.4	1257	62.8	19.8	5.8	11.5
Scientific discipline									
Medical sciences	2095	61.8	6.9	31.4	273	73.3	13.6	7.0	6.2
Humanities	1020	65.5	7.6	26.9	281	58.7	22.1	5.3	13.9
Social sciences	1039	66.1	6.8	27.0	304	53.6	28.3	4.3	13.8
Applied sciences	1489	61.7	4.4	33.8	171	73.1	12.3	7.0	7.6
Natural sciences	2307	48.8	4.1	47.1	226	60.6	19.0	6.2	14.2
Total	7950	59.0	5.7	35.3	1255	62.9	19.8	5.8	11.4
Gender									
Male	5052	56.8	5.3	37.9	890	62.4	19.1	5.7	12.8
Female	2927	62.5	6.4	31.1	367	64.0	21.5	6.0	8.4
Total	7979	58.9	5.7	35.4	1257	62.8	19.8	5.8	11.5

(1) Total numbers were not given to inhibit direct identification

Table 2. Results of a multinomial logistic regression analysis researching the composition of the post-doctoral researchers born from 1965 onwards with respect to where they obtained their PhD

	<u>Independent variable</u>	<u>b</u>	<u>se</u>	<u>Prob.</u>	<u>Odds</u>
PhD at another Flemish university	Intercept	-1.856	.154	.000	
	University of post-doc ⁽¹⁾				
	LSU1	-1.139	.144	.000	.320
	LSU2	-1.456	.150	.000	.233
	SMSU1	-.348	.168	.038	.706
	SMSU3	.770	.193	.000	2.159
	Scientific discipline ⁽²⁾				
	Medical	.529	.144	.000	1.697
	Humanities	.507	.165	.002	1.660
	Social sciences	.295	.168	.080	1.343
	Applied sciences	.251	.174	.149	1.285
Gender ⁽³⁾					
Men	-.009	.103	.927	.991	
PhD outside Flanders	Intercept	-.082	.083	.324	
	University of post-doc ⁽¹⁾				
	LSU1	-.632	.085	.000	.531
	LSU2	.142	.078	.066	1.153
	SMSU1	-.044	.102	.663	.957
	SMSU3	-.355	.167	.034	.701
	Scientific discipline ⁽²⁾				
	Medical	-.566	.066	.000	.568
	Humanities	-.793	.085	.000	.453
	Social sciences	-.800	.084	.000	.449
	Applied sciences	-.535	.072	.000	.586
Gender ⁽³⁾					
Men	.197	.052	.000	1.217	

Model Chi² = 634.569 p < .001

n = 7979

Note: The dependent variable in this analysis is the place where the PhD was obtained which is coded so that the category PhD within the same university is used as reference; (1) a middle sized university is used as reference; (2) natural sciences are used as reference; (3) females are used as reference

Table 3. Results of a multinomial logistic regression analysis researching the composition of the ZAP-staff subjects born from 1965 onwards with respect to their institutional origin

	<u>Independent variable</u>	<u>b</u>	<u>se</u>	<u>Prob.</u>	<u>Odds</u>
At least 1 switch within Flanders	Intercept	-.285	.225	.205	
	University of ZAP ⁽¹⁾				
	LSU1	-1.309	.215	.000	.270
	LSU2	-1.889	.245	.000	.151
	SMSU1	-.672	.246	.006	.511
	SMSU3	.090	.337	.789	.1094
	Scientific discipline ⁽²⁾				
	Medical	-.343	.260	.187	.709
	Humanities	.334	.240	.165	1.396
	Social sciences	.627	.230	.006	1.871
	Applied sciences	-.252	.306	.410	.777
No academic history in Flanders	Intercept	-1.365	.308	.000	
	University of ZAP ⁽¹⁾				
	LSU1	-.488	.317	.124	.614
	LSU2	.085	.302	.777	.1089
	SMSU1	-.351	.384	.361	.704
	SMSU3	.759	.452	.093	2.136
	Scientific discipline ⁽²⁾				
	Medical	-.980	.322	.002	.375
	Humanities	.066	.268	.806	1.068
	Social sciences	.131	.264	.620	1.140
	Applied sciences	-.706	.357	.048	.494
PhD outside Flanders, post-doc & ZAP in Flanders	Intercept	-1.822	.386	.000	
	University of ZAP ⁽¹⁾				
	LSU1	-1.424	.434	.001	.241
	LSU2	-.345	.369	.350	.708
	SMSU1	-.300	.440	.494	.741
	SMSU3	.418	.552	.450	.1518
	Scientific discipline ⁽²⁾				
	Medical	.009	.374	.980	1.009
	Humanities	-.043	.394	.914	.958
	Social sciences	-.199	.406	.624	.819
	Applied sciences	.163	.424	.701	1.177
Model Chi² = 165.598 p < .001					
n = 1255					

Note: The dependent variable in this analysis is the academic origin which is coded so that the category 'entire career at the same Flemish university' is used as reference; (1) a middle sized university is used as reference; (2) natural sciences are used as reference



Gatekeeping African studies: A preliminary insight on what do editorial boards indicate about the nature and structure of research brokerage

Sandro Mendonça*, João Pereira* and Manuel Ennes Ferreira**

*sfm@iscte.pt

ISCTE Business School, BRU and UECE, and SPRU (University of Sussex)

** ISEG, University of Lisbon

joao.joaopedro@gmail.com; mfereira@iseg.ulisboa.pt

ABSTRACT

Issues of inequality and distribution of different sorts constitute salient aspects in the agenda of development studies. As globalisation unfolds new indicators are needed in order to update the capacity identify, capture, and interpret its dynamics and asymmetries. Notwithstanding the many changes in the world's socio-economic landscape, Africa, in particular, remains a marginal geographical and cognitive space. This research deals with a particular segment of the development studies "academic value chain". Unlike much of the contemporary work on scientometrics, this paper does not primarily analyse publications or citations. Our data regards the composition and characteristics of editorial boards in the field under analysis.

INTRODUCTION

From the periphery globalisation all too often understood as a polarized phenomenon. Indeed, issues of inequality and distribution of different sorts have constituted salient aspects in the agenda of development studies. As globalisation unfolds new indicators are needed in order to update the capacity identify, capture, and interpret its dynamics and asymmetries. Notwithstanding the many changes in the world's socio-economic landscape, Africa, in particular, remains a marginal geographical and cognitive space. Renewed effort is needed to empirically understand the cognitive locus of this space amidst a globalising, but asymmetric, process of science, technology, innovation and broader economic change.

GOING UP THE VALUE CHAIN OF AFRICAN STUDIES

This research deals with a particular segment of the development studies "academic value chain". Here we concentrate in the academic research on social-economic development and we take as object the research outlets that themselves take development as their object. Specifically, we focus on the top journals that delve with African studies.

Journals are fundamental institutions in the contemporary sector trying to understand and build policy approaches. They offer and certificate repertoires of attitudes and policy stances for engaging with big development challenges. However, little is known about peer-reviewed international research journals and our paper hopes to provide a first look into this black-box.

APPROACH AND PRELIMINARY RESULTS

Unlike much of the contemporary work on scientometrics, this paper does not primarily analyse publications or citations. Our data regards the composition and characteristics of editorial boards in the field under analysis.

Editors play a central role in scholarly governance. These groups of scholars have executive and consulting responsibilities that ultimately select and mould the research that is academically validated, formally published and readily accessible in general. So far the profiles of these roles have not been subject to systematic scrutiny. We do this, at an exploratory level, through the prism of a number of variables: geography, gender, affiliation, etc.

Our paper focuses on the five leading journals in African studies, the older one being African Affairs, and compiles information regarding its 239 editors (as of March 2016). We show that Africa-based editors constitute a minority of total editors (just under one third). Among African South Africans are dominant, followed by Nigerians, Tanzanias and Botswanians.

Northern hemisphere Anglo-American editors comprise the larger community (130 out of 239). For instance, the two chief editors of the *Journal of Modern African Studies* are located in Britain and the US. European editors the largest overall continental contingent, 118 in total (23 of which non-British).

Female editors are 27.2% of this elite academic population, ranging from about 14% to 35% in our sample of journals. What is more, 22.7% of the African-based editors are women, whereas 30% is the case for the European ones. The *Journal of South African Studies* is the outlet with the greater gender balance.

Regarding affiliation 15.5% of the editors have non-academic affiliations. Such institutions are governmental agencies, NGOs, museums, etc. Most non-academics are located in European think-tanks. The *ROAPE* journal concentrates 47.4% of these non-academics.

IMPLICATIONS

This paper tries to cast some light into the structure of scholarly publishing. It seems to be the first to document in any systematic way how editorial groups are structured. Many challenges remain. The actual editorial decisions remain unobserved. Their discussions and the handling of papers are not publicly available. This empirical study can only draw general inferences regarding the key characteristics of boards based on publicly available data.

Given the uncertainty about how editors actually operate, researchers may strive to come with other complementary indicators along these lines (for a related approach see Adams et al., 2010, and Schwartz-Ziv and Weisbach, 2013). For instance, more research can be done regarding the different roles of editors inside the same journal. This remains work for further research and we sketch how such an agenda could look like.

REFERENCES

- Adams, R.B., B.E. Hermalin, & Weisbach, M.S. (2010), "The role of boards of directors in corporate governance: A conceptual framework and survey", *Journal of Economic Literature*, 48, 1, 58-107.
- Schwartz-Ziv, M. & Weisbach, M.S. (2013), "What do boards really do - Evidence from minutes of board meetings", *Journal of Financial Economics*, 108, 2, 349-66.



The Network of International Student Mobility (Working Title)

Eva Maria Vögtle*, Michael Windzio**

*voegtle@dzhw.eu

Department of Governance in Higher Education and Science, DZHW (German Centre for Higher Education Research and Science Studies), Lange Laube 12, 30159 Hanover (Germany)

** mwindzio@empas.uni.bremen.de

EMPAS (Institute for Empirical and Applied Sociology), University of Bremen, Mary-Somerville-Str. 9, Unicom, Haus Salzburg (Eingang 9), 28359 Bremen (Germany)

ABSTRACT

In a previous paper (see Vögtle and Windzio 2016) we investigated the impact of membership in the Bologna Process on patterns and driving forces of cross-national student mobility. Student exchange flows were analysed for almost all Bologna Process member states and non-Bologna OECD members over a ten-year period (from 2000 to 2009). We applied a *Social Network Approach* focusing on outbound diploma-mobility to identify positions of countries in the network of cross-national student exchange. Based on social network analyses, we first visualized the exchange patterns between sampled countries. In doing so, we analysed the student exchange linkages to gain descriptive insights into the development of the network (see Figure 1). Second, we used *Exponential Random Graph Models* (ERGM) to test which factors determine patterns of transnational student mobility. The results of this network analyses reveal that cross-national student exchange networks are stable over time. At the core of these networks are the United States, Great Britain, France, and Germany; they attract the highest shares of students from the remaining countries in our sample. Moreover, the results of the ERGM demonstrate that homophily between countries determines student exchange patterns. The most relevant ties exist between bordering countries. Moreover, membership in the Bologna Process impacts on mobility patterns, and the effect size increases over the periods investigated.

INTRODUCTION

In Western democracies, public universities are often regarded as role models for the democratic organisation of the overall society. In the context of international student mobility, it is often theorized that students from democratically less developed countries might cater to the diffusion of democratic norms in their home country upon return after a study related stay in a democratically more advanced country. Up to date, there is not yet a debate if international student mobility can have such an impact. In this ongoing project, our particular focus is on countries' levels of democracy according to the Democracy Index of the Intelligence Unit of The Economist (2010). Our research aims at investigating the flow of internationally mobile students from a network perspective. We expect that the higher levels of democratization are, the more attractive countries become as destinations for internationally mobile students. Due to the positive correlation between economic development and democratization, which we find in our data, we focus on the mediating effect of GDP per capita.

DRIVING FORCES OF TRANSNATIONAL NETWORK FORMATION: PREVIOUS RESULTS

In a previous article (see Vögtle and Windzio 2016) we investigated the impact of membership in the Bologna Process on patterns and driving forces of cross-national student mobility. Student exchange flows were analysed for almost all Bologna Process member states and non-Bologna OECD members over a ten-year period (from 2000 to 2009). We applied a *Social Network Approach* focusing on outbound diploma-mobility to identify positions of countries in the network of cross-national student exchange. We visualized the exchange patterns between sampled countries and analysed the student exchange linkages to gain descriptive insights into the development of the network. This revealed that cross-national student exchange network is stable over time with the United States, Great Britain, France, and Germany at its core; they attract the highest shares of students from the remaining countries in our sample.

Analytically, we refer to the *principle of homophily*, derived from social network theory (McPherson et al. 2001), as well as assumptions about the *complementarity* of exchange relationships. This already implies that they can be imbalanced and characterized by *dependence structures and power asymmetries* (see Pfeffer and Salancik 1978). We refer to theories of *hegemony*, which regard the dominance or supremacy of an institution, state, organization or similar actor as a driving factor for alliances (see Shields 2013:615). According to the results of previous studies on international student mobility (Barnett and Wu 1995, Chen and Barnett 2000, Vögtle and Windzio 2016), an *academic hegemony* (Barnett and Wu 1995) consistent with world economic and political performance can be assumed.

Through previous research we have identified that a) common membership in the Bologna-Process leads to a higher share of international exchange students, that b) the most relevant but also most imbalanced (see Vögtle and Fulge 2013; Fulge and Vögtle 2014) exchange relationships exists between bordering countries, that c) English speaking countries attract the highest share of international students and that d) countries with a high GDP per capita have greater chances to host students than less prosperous countries. We now wish to widen our focus by including variables accounting for the degree and similarity in democratization level between the sampled countries. We wish to investigate the democratization aspects in more detail, underpin it theoretically better and include a broader data base for the assessment of a countries' level of democracy.

DATA, SAMPLE AND METHODS

The data on internationally mobile students reflects absolute numbers of dyadically exchanged students in the respective years and can be conceptualized in the form of a series of directed links. We analyze the number of exchanged students divided by the total number of students enrolled in tertiary education (International Standard Classification of Education [ISCED 1997] level 5A and 6) in the sending country which accounts for the size of the tertiary education sector of the countries investigated. Thus, we account for the opportunity of the whole student body of a given country to become internationally mobile.

The dependent variable refers to the number of students EGO (sending country) who migrate to ALTER (receiving country), divided by the number of all students enrolled in EGO. After computing the quartiles over all country-pairs, the dependent variable has been dichotomized by setting the highest quartile to one and the three lower quartiles to zero for the analysis of

binary networks (see models 1-3 in Table 2). By imposing this restriction to the definition of a network tie in the respective dimension, we ensure that only a relevant share of ego's students who migrate to the alter country constitute a tie between the two countries. As has been outlined above, *homophily* assumptions are central to our investigation. We thus analyze whether, on the one hand, (a) structural similarity in higher education policies between countries (e.g. the common use of ECTS), (b) general socio-economic similarities or (c) cultural similarities influence the probability for exchange ties between the countries under investigation. Previous studies on cross-national student mobility found out that the flow of international students is closely tied to economic development (see Barnett and Wu 1995, Chen and Barnett 2000), thus, we use general socio-economic similarities—in our case measured as the (negative) absolute distance between the GDP per-capita of two countries—to account for national capabilities. According to a recent study on international student mobility, a common language between two countries is among the most important predictors for the structure of the flow of exchanged students (see Barnett et al. 2015), we thus account for cultural similarity by coding whether two countries' official language(s) belong to the same language family or language branch (see Lewis et.al. 2015). Additionally, since the existence of a common border between two countries and geographic proximity between them have been identified as important predictors for student exchange patterns between them (see Barnett et al. 2015), we include information as to whether two countries of a dyad share a common land or sea border. Secondly, we use exponential random graph models (ERGMs) that were developed in order to explain why we observe a specific empirical realization of a network among a given set of units (vertices). Here, the focus is on the structural characteristics of the network, on the attributes of countries as well as on their similarity or difference with regard to salient characteristics, which determine whether there is a tie between them or not. We estimate ERGMs (Lusher et al. 2013) for binary networks, where the dependent variable is 1 if there is a relevant tie in a dyad, and 0 otherwise.

DETERMINANTS OF TIES: RESULTS OF ERGM

Table 1 shows a series of exponential random graph models of ties in the network of global student mobility for the years 2000, 2004, and 2009. For each year we estimated three models: the first model includes the effects of levels of democratisation for in-degree and out-degree as well as the effect of the absolute difference in democratisation. In addition, network structural effects have been controlled, namely GWESP (geometrically weighted edgewise shared partners), mutuality and cyclic triplets. There is also a dyadic edge-covariate indicating whether two countries share a border or not. The second model controls also for the effect of languages, namely of English on in-degree, the same language branch, as well as a mix of languages in the focal country on in- and out-degree. In addition, the second model controls for ECTS and membership in the Bologna process. In the third model specification, finally, we also control for GDP per capita on in-degree, out-degree, and the effect of the absolute difference in GDP on a ties in the respective dyad.

In models 1 and 2 we find a negative effect of the absolute difference in democratisation. In other words, the more dissimilar two countries are with respect to democratization, the lower is the propensity to form a tie in the network of student mobility. Possibly, student mobility is concentrated at the two ends of the continuum of democratisation: there could be a high density within comparatively authoritarian regimes on the one hand, but also within comparatively democratic regimes. We will readdress this issue below. In addition, in models 1 and 2 we find highly significant and positive effects of democratisation on in-degree in

2000. In the subsequent years the effect seems to decline and loses in statistical significance, albeit it is still significant at the 5% level in 2009.

The most striking result is the change in the effect of democratisation on in-degree in models 3, 6 and 9, after controlling for GDP per capita. With regard to the GDP effects, we observe an increase in GDP per capita in 2000, in-degree as well as the out-degree in the network increases. The absolute difference in GDP is significant and positive only in the year 2009 (model 9). In other words, countries with high GDP per capita sent considerable proportions of their student population to many other countries only in 2000, but have received students from many other countries in all years. Wealthy countries seem to be highly attractive for students from many other countries. However, in combination with the positive effect of in-degree and the insignificant effect of out-degree, the positive effect of the absolute difference in GDP per capita in 2009, we find indeed that many less wealthy countries send students to comparatively wealthy countries.



Table 1. Ties in the network of international student mobility 2000-2009, ERGMs

	2000		network tie 2004					2009	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
edges	-6.577***	-6.231***	-1.757*	-4.627***	-5.261***	-1.536*	-4.145***	-5.368***	-1.858*
mutuality	-0.030	-0.229	-0.298	0.696**	-0.021	-0.021	0.340	-0.466	-0.514
GWESP (alpha=0.693)	1.541***	1.487***	1.126***	1.639***	1.513***	1.242***	1.713***	1.652***	1.347***
cyclic triplets	-0.551***	-0.480***	-0.491***	-0.654***	-0.488***	-0.464***	-0.759***	-0.637***	-0.565***
economic performance									
absolute difference(GDP per capita)	-	-	-0.013	-	-	0.005	-	-	0.036***
GDP per capita: indegree	-	-	0.133***	-	-	0.123***	-	-	0.093***
GDP per capita: outdegree	-	-	0.050***	-	-	0.024	-	-	0.019
democratization									
authoritarian regime: indegree	0.292***	0.213***	-0.424***	0.143**	0.140**	-0.538***	0.080*	0.111*	-0.489***
authoritarian regime: outdegree	0.087	0.043	-0.156	-0.010	-0.018	-0.120	-0.030	0.013	-0.190*
absolute difference (democracy index)	-0.233***	-0.188**	-0.343***	-0.308***	-0.243***	-0.441***	-0.243***	-0.188***	-0.556***
hegemonic language									
english language: indegree	-	0.693***	0.307*	-	1.114***	0.698***	-	0.961***	0.911***
same language branch	-	0.599***	0.636***	-	0.484**	0.565***	-	0.322*	0.532**
mixed language: indegree	-	0.335	-0.254	-	-0.025	-0.606**	-	-0.226	-0.881***
mixed language: outdegree	-	0.259	0.062	-	0.294	0.190	-	0.086	-0.031
geographic proximity									
shared border	2.283***	2.066***	2.231***	-	2.071***	2.202***	-	2.061***	2.091***
Bologna process									
both apply ECTS or comparable system	-	0.110	0.087	-	-0.026	-0.262	-	-0.192	-0.070
both countries are Bologna participants	-	0.532***	0.534***	-	0.534***	0.739***	-	0.689***	0.834***
Akaike Information Criteria	1,410.055	1,360.051	1,238.573	1,488.969	1,320.952	1,203.299	1,451.708	1,310.553	1,190.470
Bayesian Information Criteria	1,453.274	1,435.685	1,330.415	1,526.786	1,396.587	1,295.141	1,489.525	1,386.187	1,282.312

Note:

*p<0.05; **p<0.01; ***p<0.001



Interestingly, there is an inter-correlation of GDP per capita and democratisation, which reverses the effect of democratisation on in-degree: in all three years the effects of democratisation remain robust after controlling for the effects of language, ECTS and membership in the Bologna process in the first two models for each year. Not until the effects of GDP per capita are controlled in each third model, the effect of democratisation on in-degree reverses its sign. In other words, conditional on levels of economic development, high levels of democratisation show a negative effect on in-degree. Highly democratic countries are unattractive as destinations of student mobility if they're not at the same time economically highly developed.

In addition to the edge effect that is the intercept of the ERGM regression, we determine four characteristics related to the embeddedness of a country pair or dyad in the social network. First, in the binary models we have no significant and weak effects of mutuality, indicating that if one country sends a high share of students to another country, it is not reciprocated to a similar extent. Effects of the GWESP (geometrically weighted edgewise shared partners) are strong and significant: the positive effect of GWESP indicates high degrees of transitivity. Taken together, there is a strong tendency to triadic closure. The presence of each triangle involved increases the likelihood of attaching further triangles with a decreasing function of the number of triangles (Lusher et al. 2013:71). Moreover, the effect of cyclic triangles tends to be negative. Cyclic triangles indicate a non-hierarchical circuit exchange, which is not common between countries, according to the result. Rather, ties between countries due to mobility patterns tend to be hierarchical (see Lusher et al. 2013:44).

SUMMARY AND FURTHER RESEARCH

With regard to homophily assumptions, we can state that an important factor positively influencing student exchange patterns over all models estimated is geographical propinquity (see Shields 2013), which is in line with macro-level theories of migration. Another important homophily factor is the common usage of languages of the same language branch. Unlike common Bologna membership, common application of the ECTS or a comparable system does not increase the likelihood of an exchange tie between two countries. Cultural hegemony, singling out English-speaking countries, is influential. There exists an inter-correlation of GDP per capita and democratisation, highly democratic countries are unattractive as destinations of student mobility if they're not at the same time economically highly developed. In addition, we see that the basic pattern of hierarchical triadic relationships holds for the countries' student mobility networks. Hence, the networks of international student mobility investigated resemble patterns commonly observed in friendship networks.

In further research, we wish to include non-OECD countries with high outbound-mobility rates into our sample (e.g. China, India, and Pakistan). We seek to test if our results are still valid if the investigation moves beyond OECD countries and if different mechanisms determine the flow of international student mobility if our sample is extended to include these large international student export nations. Moreover, this will allow us to compare different networks, the network of mere OECD countries with the network when large Asian countries are included in the sample. Moreover, we will investigate if religious aspects impact on international student mobility flows, thus we wish to extend cultural similarity variables already investigated (see Vögtle and Windzio 2016) by variables covering religious similarities between countries. This might be even more interesting when we further extend

our sample in order to include countries with predominantly Muslim population. Moreover, we wish to incorporate analyses based on valued relations to model structural network effects.

References

- Barnett, G.A., Lee, M., Jiang, K., & Park, H.W. (2015) The Flow of International Students from a Macro Perspective: A network Analysis. *Compare: A Journal of Comparative and International Education*, doi:10.1080/03057925.2015.1015965.
- Barnett, G.A., Wu, R.Y. (1995): The International Student Exchange Network: 1970&1989. *Higher Education*, 30, 353-368.
- Chen, T., & Barnett, G.A. (2000). Research on international student flows from a macro perspective: A network analysis of 1985, 1989 and 1995. *Higher Education*, 39, 435-453.
- Fulge, T., & Vögtle, E. (2014). Sweeping Change - But Does It Matter? The BP and Determinants of Student Mobility. In Martens, K., Knodel, P., & Windzio, M. (Hrsg.) *A New Constellation of Statehood in Education Policy?*. Basingstoke: Palgrave Macmillan: 67-88.
- Lewis, M. P., Simons, G.F. & Fennig, C.D. (eds.) (2015). *Ethnologue: Languages of the World, Eighteenth edition*. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>. Accessed 13th July 2016.
- Lusher, D., Koskinen, J., & Robbins, G. (2013). *Exponential random graph models for social networks*. Theories, methods, and applications. Cambridge: Cambridge University Press.
- McPherson, J. Miller, Lynn Smith-Lowin and James M. Cook. 2001. Birds of a feather: Homophily in social networks. In: *Annual Review of Sociology* 27, 415–444.
- Pfeffer, Jeffrey and Gerald A. Salancik. 1978. *The External Control of Organisations - A Resource Dependence Perspective*: New York: Harper Row Publishers.
- Shields, R. (2013). Globalization and International Student Mobility: A Network Analysis. *Comparative Education Review*, 57(4), 609-636.
- The Economist Intelligence Unit, The Democracy Index 2010: Democracy in retreat. http://www.eiu.com/public/topical_report.aspx?campaignid=demo2010 Accessed 13th July 2016.
- Vögtle, E., & Fulge, T. (2013). *The Bologna Process' impact on cross-national student mobility. A multi-method evaluation*. TranState Working Papers 179. SFB Staatlichkeit im Wandel, Universität Bremen.
- Vögtle, E.M. & Windzio, M. (2016). Networks of International Student Mobility: Enlargement and Consolidation of the European Transnational Education Space? *Higher Education*, DOI 10.1007/s10734-015-9972-9



Scientific mobility of Early Career Researchers in Spain and The Netherlands through their publications¹

Nicolas Robinson-Garcia^{*}, Carolina Cañibano^{*}, Richard Woolley^{*} and Rodrigo Costas^{**}

^{*}elrobin@ingenio.upv.es; cacaisan@upvnet.upv.es; ricwoo@ingenio.upv.es
INGENIO (CSIC-UPV), Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia (Spain)

^{**}rcostas@cwts.leidenuniv.nl
Centre for Science and Technology Studies (CWTS), Leiden University, Leiden (The Netherlands)

ABSTRACT

This paper presents an exploratory analysis of different typologies of researchers according to their traceable mobility using scientific publications covered in the Web of Science (WoS). We compare two populations of researchers, of the same 'scientific age', based in Spain and The Netherlands. To establish reasonable comparisons between researchers based in The Netherlands and Spain, we must first identify similar groups of researchers in each country. We only consider 'trusted' direct linkages between author and affiliation as reported in scientific publications and recorded by WoS. We establish three different study groups: Mobile versus non-mobile researchers, returned versus not returned researchers and, single versus multiple affiliations. We observe differences in the mobility patterns and their relation with production and citation impact between countries. Differences for each study group are found in the case of Spain but not as evident for The Netherlands. We conclude remarking the need to further analyse the institutional framework of each country to better understand how much do they influence research mobility and in what way.

INTRODUCTION

International scientific mobility is acknowledged to be a key mechanism for the diffusion of knowledge, particularly tacit or 'sticky' knowledge that cannot be transferred without geographical proximity and personal contact (Frenken 2010), for the incorporation of young researchers into elite transnational scientific networks (Laudel 2005), and for accessing additional resources or infrastructures that are essential to the research process but located in other places. The benefits of mobility as a means for knowledge diffusion are used as justification for the support policies and programmes that encourage the mobility of European Researchers (CEC 2004, 2008). The demand for researcher mobility indicators has been building in the EU since the launching of the Mobility Strategy for the European Research Area (CEC 2001).

The inadequacy and lack of appropriate data to assess the phenomenon of researcher mobility has been repeatedly pointed out by scholars and policy makers (CEC 2004; Fontes 2007; Didou-Aupetit 2009; Ackers 2009). Traditional comparable data sources such as migration or labour statistics and censuses allow us to picture flows of highly skilled human resources

¹ Nicolas Robinson-Garcia is currently supported by a Juan de la Cierva-Formación Fellowship from the Spanish Ministry of Economy and Competitiveness

(HRST) (Cañibano & Woolley 2015) but not the movements of researchers which is more complex and often circulatory (Jöns 2007). However, important progress in the understanding of the phenomenon has been made building on the direct collection of information from researchers via surveys (IDEA Consult 2013), interviews (Ackers 2009) and CV analysis (Cañibano et al, 2011).

From a bibliometric perspective, Laudel (2003) suggests using the address field of publications to identify mobility patterns between researchers. While this approach seems feasible for studies based on small samples, it has been problematic at a large scale as researchers must be identified univocally. Author disambiguation is a historically problematic issue in bibliometric research when applied to individual scholars (Smalheiser & Torvik, 2009). One of the major limitations has been the lack of reliable methods for the proper identification (i.e. overcoming the usual problems of homonyms and synonyms found in personal names) of the full scientific output of the authors (Costas & Bordons, 2009; Reijnhoudt et al., 2014).

In recent years, two different methods have been proposed for dealing with this problem. One is the Author ID developed by Scopus; the other is the algorithm developed by Caron and van Eck (2014). The development of such methodologies, together with the incorporation of the linkages between authors and affiliations in the records covered in two of the most important bibliographic databases (Web of Science and Scopus) has allowed the exploration of individual address changes in publications at a large scale. In this regard, Moed and colleagues have worked on the potential of bibliometric data as a proxy to measure mobility using the Scopus Author ID (Moed, Aisati, & Plume, 2013; Moed & Halevi, 2014),

This paper builds on their work and presents an exploratory analysis of different typologies of researchers according to their traceable mobility using scientific publications covered in the Web of Science (WoS). We compare two populations of researchers, of the same ‘scientific age’, based in Spain and The Netherlands. We address the following research questions:

1. Are there substantial differences in terms of research publication and citation impact between mobile researchers and non-mobile researchers within each country?
2. When does mobility take place in the research trajectory of a researcher? For those who return to their original country, how many years does it take?
3. Are there differences between the mobility profiles, publication productivity and citations of the Spain-based and Netherlands-based researchers?

The next section details the data and methods of analysis used and the following section specifies the results. The final section includes discussion of the results and of contextual factors associated with the Spanish and Dutch research systems that may contribute to explaining the initial outcomes we observe.

MATERIAL AND METHODS

To establish reasonable comparisons between researchers based in The Netherlands and Spain, we must first identify similar groups of researchers in each country. We only consider ‘trusted’ direct linkages between author and affiliation as reported in scientific publications and recorded by WoS. In other words, we don’t expect false positives in our linkages (i.e. wrong linkages between authors and affiliations) although we may have false negatives (i.e.

we may miss linkages between the authors and the affiliations that have not been recorded in the publications). We define researchers' scientific age based on the year when they published their first reported paper with a target affiliation. As we aim at analysing career trajectories, we focus on researchers who published their first paper between 2003 and 2005. This way we ensure that they are all of a similar 'generation' (Costas, Nane & Larivière, 2015). We then define their home country as the one to which they were affiliated in their first year of publication. We work only with researchers who published their first paper whilst affiliated to an organization in either Spain or The Netherlands. Based on this set of researchers, we build their full publication record up to 2014. That way we can analyse their full publication career. We establish three different study groups:

Mobile versus non-mobile researchers. We consider a researcher as mobile when at any point in time during their career they have declared more than one country in the address field of one or more of their publications.

Returned versus not returned researchers. We consider a researcher has returned to their home country when they have reported within their later publications that they are again affiliated within the same country in which they published their first paper, following a period in which they were affiliated in a different country. This does not mean that it is the actual 'home' country of the researcher, rather it is their country of scientific origin as derived from their publication profile. In cases of double affiliation in The Netherlands and Spain, we have considered them as belonging to both countries. That is, a small sample of our data is duplicated for each country.

Single versus multiple affiliation. One of the limitations we found in our study is the role of researchers who report an affiliation to more than one institution or country. For that reason, we decided to analyse these special cases as a separate group.

These three study groups are not distinct from one another. In fact, the two latter are overlapped subgroups of the mobile researchers group. At this point we did not cross different variables as we wanted to conduct a first analysis of the dichotomous distinction between the groups and make comparisons between the two countries.

RESULTS

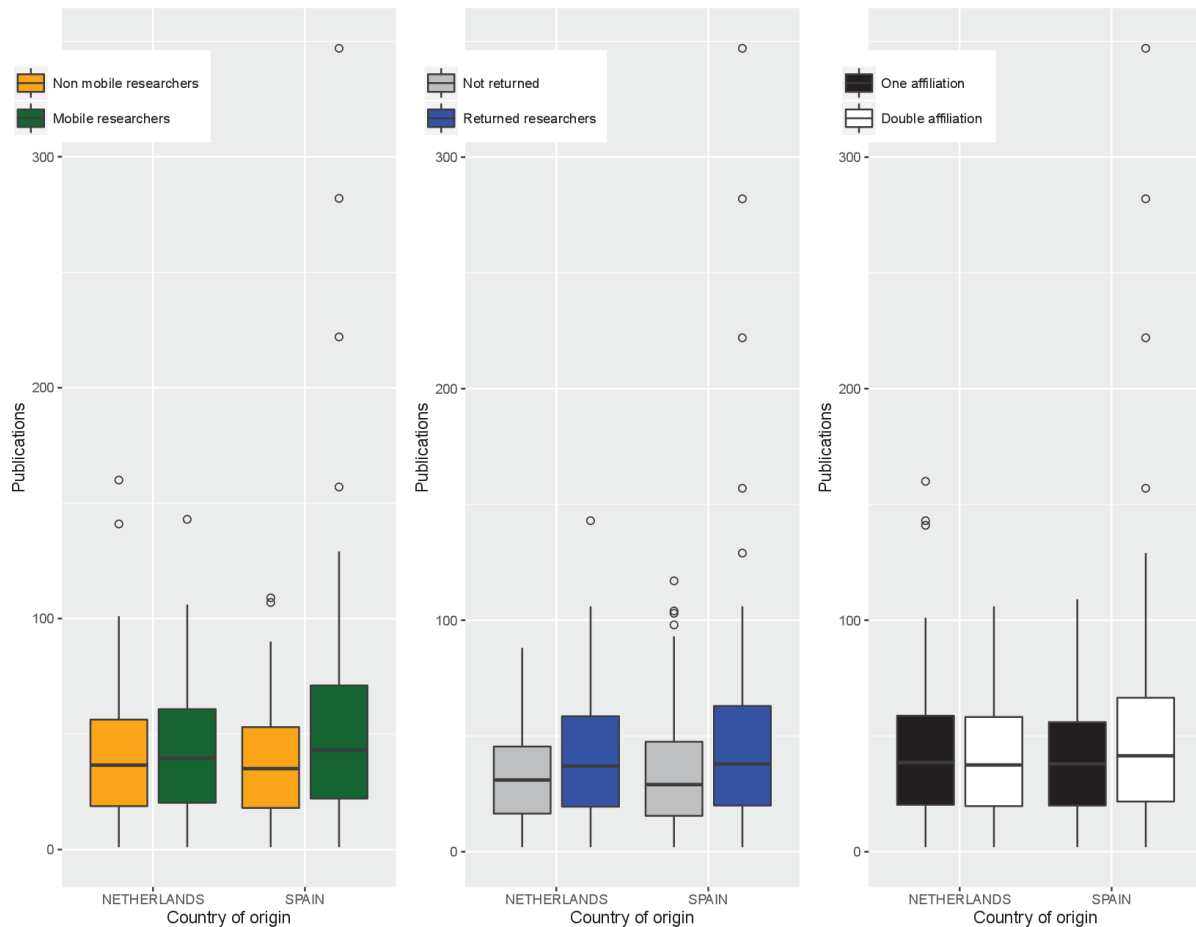
Our data set includes a total of 10,412 researchers. Table 1 shows their distribution among the three classes of study groups. More than 20% of researchers from The Netherlands reported some degree of international mobility. Spain shows a lower international mobility rate (13.4%). However, while only 39.7% of mobile researchers from The Netherlands (n=344) had 'returned', the share increases up to 60.8% of mobile researchers from Spain (n=501). With regard to multiple affiliations, the share of researchers who reported being affiliated to institutions in more than one country at the same time is higher in the case of The Netherlands (9.81% of the total share) than for Spain (7.35%).

Table 1. Descriptive indicators of the study groups and home country

	SPAIN		THE NETHERLANDS	
	N	%	N	%
Non-mobile researchers	5327	86.60	3394	79.65
Mobile researchers	824	13.40	867	20.35
Returned researchers	501	8.15	344	8.07
Multiple affiliation	452	7.35	418	9.81
Total	6151	100.00	4261	100.00

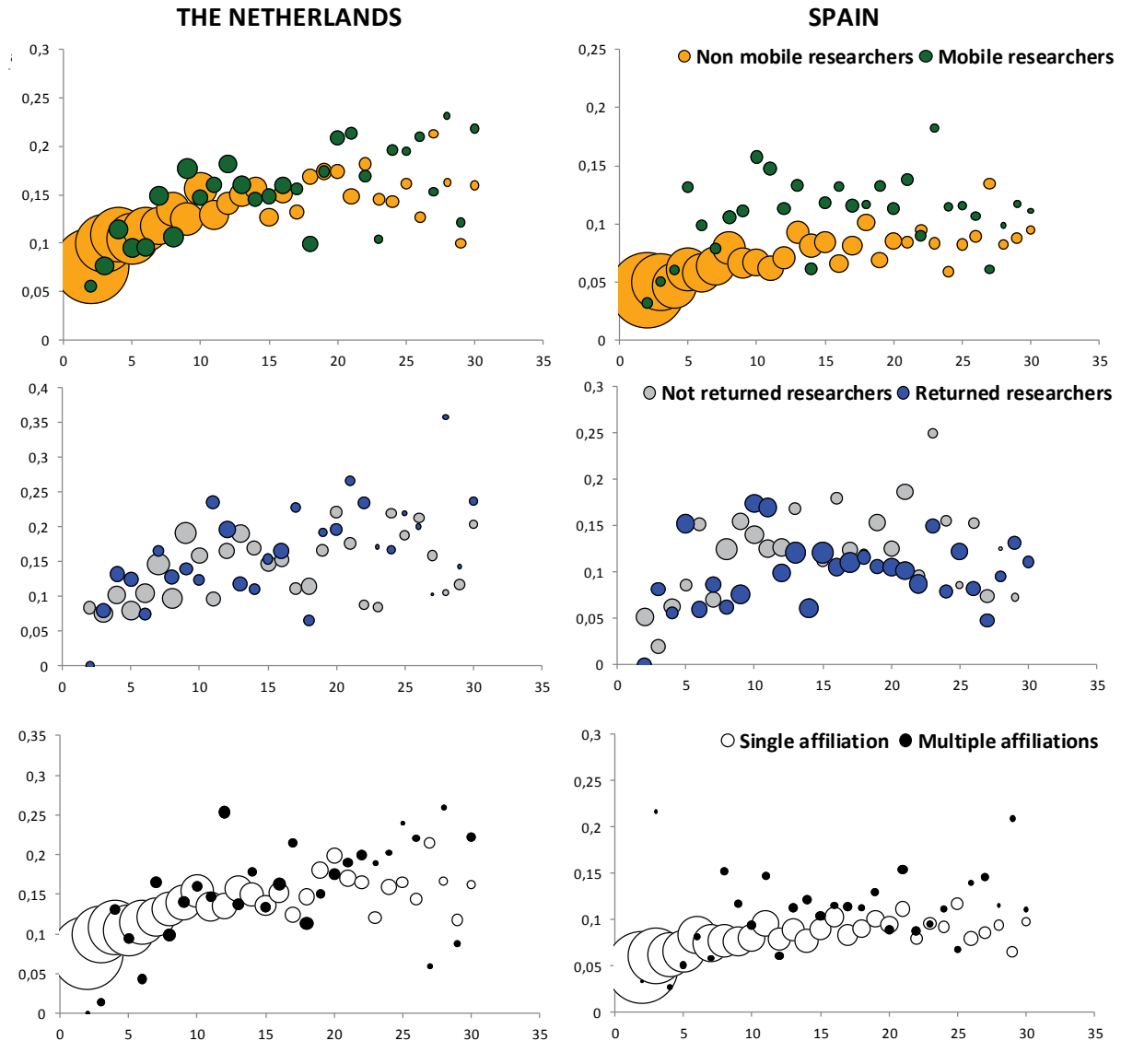
Figure 1 compares research productivity in terms of publications for the three classes of study groups and for The Netherlands and Spain. The boxplot on the left compares non-mobile with mobile researchers. Mobile researchers tend to have higher levels of production than non-mobile researchers. Differences are very small in the case of The Netherlands but slightly more pronounced for Spain. The middle boxplot compares returned vs. not-returned researchers. Here we find larger differences between the two groups, both for Spain and The Netherlands. Returned researchers show higher levels of publications production than those mobile researchers who have not returned home. Finally, at the right is the comparison between researchers with a single affiliation and those with multiple affiliations. An almost identical distribution is apparent in the case of The Netherlands and a slight difference in distribution is evident for Spain, although the average number of publications per researcher in Spain is very similar regardless of the number of affiliations.

Figure 1. Boxplot on publications distribution by study group and home country



With regard to the relationship between the type of researcher and citation impact, we relate the share of highly cited papers in comparison with the total number of publications. Figure 2 compares each researcher type by country. In the case of the Netherlands, there seems to be no difference in the citation impact of researchers depending on their mobility and affiliation attributes. (However it should be noted that data is sparser for returned versus not-returned researchers and for researchers with multiple affiliations.)

Figure 2. 10% Highly Cited Papers and publications by study group and home country. Size represents number of researchers. Y axis represents share of highly cited papers, X axis represents number of publications. Only included, researchers with >2 and <31 publications



In contrast, we do find differences in the Spanish case for the three classes of groups, showing higher impact for mobile, not-returned and multiple affiliation researchers. Again, we find that the relation between publication output and citation impact is not as correlated in the case of returned, not-returned and double affiliation researchers.

Table 2 analyses two different temporal aspects: number of years that it takes for mobile researchers to leave their home country; and number of years reported abroad for returned researchers. Here again we see differences between the Spain and The Netherlands.

Researchers assigned to Spain take a longer period until they leave their home country (4.6 years on average compared to 4.2 years for researchers assigned to The Netherlands) and also stay abroad for a shorter period (around 3 years on average compared to around 4 years for The Netherlands).

Table 2. Years to emigration for mobile researchers and years to return for returned researchers

NETHERLANDS	Years to emigration	Years abroad	SPAIN	Years to emigration	Years abroad
Average	4.17	3.80	Average	4.61	2.98
Std. Dev.	2.51	3.09	Std. Dev.	2.51	2.77
Population	867	344	Population	824	501

DISCUSSION AND FURTHER RESEARCH

We observe differences in the degree of mobility of Spain and Netherlands based researchers. The results are based on evidence of mobility that is long enough to lead to a change in affiliation, such as a job change or extended research visits. This does not necessarily reflect the overall degree of mobility of researchers based in Spain and the Netherlands that would be evident should other forms of mobility, such as short-term stays, also be considered.

In terms of explanations for the result, if taken at face value, factors associated with the different institutional conditions characterizing the two national systems need to be taken into account. First, the Spanish and Dutch university and research systems are different in many ways (Deloitte 2014a, 2014b; IDEA Consult 2012a, 2012b). For example, the Netherlands higher education system appears to be more internationalized than that of Spain. A higher proportion of PhD students enrolled in universities in the Netherlands are international students, compared to the situation in Spanish universities (OECD 2015). Data for 2013 show that 38% of doctoral students in the Netherlands were international students, compared to 16% in Spain and an OECD average of 24%. At the Masters level, 17% of students in the Netherlands were international students, compared to 5% in Spain and an OECD average of 14% (OECD 2015: 354). Overall, the Netherlands was hosting 2% of the global population of international students and Spain 1% (OECD 2015).

If we assume that many early career researchers will publish their first paper either during their PhD studies or shortly thereafter, then this first paper is very likely to be affiliated with the training university. In the Dutch case then, a higher proportion of these papers will be published by international students/researchers. If we then also assume that a substantial proportion of these international students will obtain a position in their home country, then the proportion of the ‘outward’ mobility from the Netherlands that is also in effect return migration by international students, will likely be higher than in the case of Spain. In other words, the more internationalized nature of the Dutch higher education system means that a higher proportion of the initial mobility included in this study is actually already part of a chain of moves. Disentangling the extent to which initial career mobility outward mobility of nationals from the training country or returning internationals is a topic for future research. At this stage we should be very cautious to not interpret our initial results as reflecting particularly Dutch or Spanish propensities for being mobile researchers. However, what seems likely is that the degree of internationalization that characterizes the research systems of the Netherlands and Spain may well include different levels of specific forms of scientific mobility, which have a different role in the research careers of the researchers concerned.

Second, there may be very different institutional incentives for mobility in the two systems. Cruz-Castro and Sanz-Menendez (2010) studied the interrelations between research mobility, productivity performance and tenure in Spain. They found that there was no return to careers from mobility in terms of accelerated progress to tenure. Due to labour market and other institutional factors associated with employment in Spain, researchers who stay within their PhD department may be advantaged through increased productivity and lower transaction costs associated with trying to access a permanent position. It is evident that mobility is not necessarily beneficial to academic careers in all respects. In fact, mobility may even be associated with a delay in tenure in the Spanish case (Sanz-Menendez et al. 2013). Whilst mobility may still have considerable value for Spain-based researchers, this may be primarily associated with internal evaluations and promotion processes, thereby creating incentives for types of mobility that do not interrupt job access and continuity within a Spanish institution. It is not directly clear how this compares to the case of the Netherlands – although the different institutional conditions that exist there (Deloitte 2014a; IDEA Consult 2012a) suggests the outcome of mobility, performance and tenure processes will produce different results to those observed in Spain. This is an area for further national and comparative research.

Another area of potential future research concerns the prevalence of researchers with multiple affiliations in the sample groups in both the Netherlands and Spain. One in ten researchers based in the Netherlands has a multiple affiliation. This *may* be linked to the internationalization of the Dutch system. However, at this stage there is no clear evidence to support this assumption. In the Spanish case the rate of multiple affiliations is only slightly lower. This may be linked to the recent wave of ‘forced emigration’ of researchers from Spain due to the severity of the financial crisis and its impact on research funding, coupled with these researchers’ desire to maintain links with ‘home’, and potential ‘return’, organisations. However, evidence is lacking and these comments are largely speculative. More sophisticated bibliometric analyses and comparisons with different ‘generation’ of researchers, possibly combined with qualitative investigation, will be required to better understand the role and function of multiple affiliations in both research mobility and research careers.

REFERENCES

- Ackers, L. (2009). *Moving people and knowledge: Scientific mobility in an enlarging European Union*. Cheltenham: Edward Elgar.
- Cañibano, C. & Woolley, R. (2015) Toward a socio-economics of the brain-drain and distributed human capital. *International Migration*, Vol. 53 (1): 115-130.
- Cañibano, C.; Otamendi, J. & Solís, F. (2011) International temporary mobility of researchers: a cross-discipline study. *Scientometrics*, Vol. 89: 653-675.
- Caron, E., & van Eck, N. J. (2014). Large scale author name disambiguation using rule-based scoring and clustering. In *19th International Conference on Science and Technology Indicators*. ‘Context counts: Pathways to master big data and little data’ (pp. 79–86).
- CEC (2001) A mobility strategy for the European research area. Communication from the commission to the council and the European parliament (COM(2001) 331 final). Brussels: CEC.

- CEC (2004) Commission staff working paper. Second implementation report on “A mobility strategy for the European Research Area” (SEC(2004)412). Brussels: Commission of the European Communities.
- CEC (2008) Realising a single labour market for researchers. Report of the ERA expert group. EUR23321EN.
- Comisión Europea (2001) Horizonte 2020, Programa Marco de Investigación e Innovación, Bruselas, 30.11.2011 COM(2011) 808 final
- Costas, R., & Bordons, M. (2009). Algoritmos para solventar la falta de normalización de nombres de autor en los estudios bibliométricos. *Investigación Bibliotecológica*, 21(42), 13-32.
- Costas, R., Nane, T., & Larivière, V. (2015). Is the year of first publication a good proxy of scholars’ academic age?. In: Ali Salah, Y. Tonta, A.A.A. Salah, C. Sugimoto & U. Al. (eds). *Proceedings of ISSI 2015 Istanbul*, pp. 988-998.
- Cruz-Castro, L. & Sanz-Menendez, L. (2010) Mobility versus Job Stability: Assessing Tenure and Productivity Outcomes, *Research Policy* 39: 27-38.
- Deloitte (2014a) Researchers’ Report 2014 Country Profile: Netherlands. Deloitte, Brussels.
- Deloitte (2014b) Researchers’ Report 2014 Country Profile: Spain. Deloitte, Brussels.
- Frenken, K. (2010) Geography of Scientific Knowledge: a proximity approach. Eindhoven Centre for Innovation Studies, Working paper 10.01
- Fontes, M. (2007) Scientific mobility policies: How Portuguese scientists envisage the return home, *Science and Public Policy* 34(4): 284–298
- IDEA Consult (2013) MORE2 Final Report. Support for Continued Data Collection and Analysis concerning Mobility Patterns and Career Paths of Researchers. Deliverable 8 prepared for the European Commission. Brussels, August 2013.
- IDEA Consult (2012a) MORE2 Country profile – Remuneration Netherlands, Idea Consult, Brussels.
- IDEA Consult (2012b) MORE2 Country profile – Remuneration Spain, Idea Consult, Brussels.
- Jöns, H. (2007) Transnational mobility and the spaces of knowledge production: a comparison of global patterns, motivations and collaborations in different academic fields, *Social Geography* 2: 97–114
- Laudel, G. (2005) Migration currents among the scientific elite *Minerva* 43: 377-395
- Laudel, G. (2003) Studying the brain drain: Can bibliometric methods help? *Scientometrics*, 57(2), 215–237.

- Moed, H. F., Aisati, M., & Plume, A. (2013). Studying scientific migration in Scopus. *Scientometrics*, 94(3), 929–942.
- Moed, H. F., & Halevi, G. (2014). A bibliometric approach to tracking international scientific migration. *Scientometrics*, 1–15.
- Reijnhoudt, L., Costas, R., Noyons, E., Börner, K., & Scharnhorst, A. (2014). ‘Seed+ expand’: a general methodology for detecting publication oeuvres of individual researchers. *Scientometrics*, 101(2), 1403-1417.
- Sanz-Menendez, L., Cruz-Castro, L. & Kenedy, A. (2013) Time to Tenure in Spanish Universities: An Event History Analysis, *PLoS ONE* 8(10): 1-18.
- Smalheiser, N. R., & Torvik, V. I. (2009). Author name disambiguation. *Annual Review of Information Science and Technology*, 43(1), 1–43.



Big Science, co-publication and collaboration: getting to the core

Michael J. Kahn*

* mikejkahn@gmail.com

University of the Western Cape, Robert Sobukwe Drive, Bellville, Cape Town 7535, South Africa

ABSTRACT

International collaboration in science has risen considerably in the last two decades (UNESCO, 2010). In the same period Big Science collaborations have proliferated in physics, astronomy, astrophysics, and medicine. Publications that use Big Science data draw on the expertise of those who design and build the equipment and software, as well as the scientific community. Over time a set of ‘rules of use’ has emerged that protects their intellectual property but that may have the unintended consequence of enhancing co-publication counts. This in turn distorts the use of co-publication data as a proxy for collaboration.

The distorting effects are illustrated by means of a case study of the BRICS countries that recently issued a declaration on scientific and technological cooperation with specific fields allocated to each country.

It is found that with a single exception the dominant research areas of collaboration are different to individual country specializations. The disjuncture between such ‘collaboration’ and the intent of the declaration raises questions of import to science policy, for the BRICS in particular and the measurement of scientific collaboration more generally.

INTRODUCTION

International collaboration in science has risen considerably in the last two decades, (UNESCO, 2010). In the same period Big Science collaborations have proliferated in physics, astronomy, astrophysics, and medicine. Publications that use Big Science data draw on the expertise of those who design and build the equipment and software, as well as the scientific community. Over time a set of ‘rules of use’ has emerged that protects their intellectual property but that may have the unintended consequence of enhancing co-publication counts. This in turn distorts the use of co-publication data as a proxy for collaboration. Such distortion may be reduced through the use of fractional counting (see e.g. NSB, 2016) but this approach is hampered by the heavy workload that is entailed in the process of allocation by institution and country.

The distorting effects are illustrated by means of a case study of the BRICS countries that recently issued a declaration on scientific and technological cooperation with specific fields of activity allocated to each country.

It is found that with a single exception the dominant research areas of collaboration are different to individual country specializations. The disjuncture between such ‘collaboration’ and the intent of the declaration raises questions of import to science policy, for the BRICS in particular and the measurement of scientific collaboration more generally.

LITERATURE REVIEW

The usual approach to measuring scientific collaboration is by the occurrence of co-authored journal publications. The quantification of co-publication has long been recognized as problematic (Katz and Martin, 1997). Part of the difficulty lies in deciding what co-publication actually means, as well as acknowledging that co-publication is prone to both double and over-counting.

Regarding international co-authorship among the BRICS countries, Kumar and Asheulova (2011) show that Russian scientists exhibit a low propensity for international co-authorship; the country also displays the highest concentration in physics and astronomy among the four BRICs. Yang et al (2012) echo Kumar and Asheulova (2011) except that both Russia and China are now shown to be highly heterogeneous across subject areas, with India mid way between the more homogeneous Brazil and South Africa. Yi et al (2013) show that South Africa is the least specialized and Russia the most, in line with Kumar and Asheulova (2011). Finardi (2015) finds that collaboration among BRICS pairs is weak compared with their collaborations with other non-BRICS countries namely the USA, United Kingdom, Germany and France. In seeking drivers of collaboration, he finds that geographic distance does not appear to have a strong bearing. In fact Waltman et al (2011) show that the mean distance between parties collaborating in science rose from 334 kilometres in 1980 to 1553 in 2009, attesting to the dispersed nature of the ‘new invisible college of science’ (Wagner, 2008).

METHODOLOGY AND DATA

Caution suggests that a study of BRICS collaboration might be better restricted to a shorter time period than is cited in the literature, say from the mid 2000s to the present. Accordingly this paper considers data for the years 2009 to 2014. Without exception the country address of each contributing author serves as the search keyword to seek co-publication counts across countries.

For completeness Table 1 summarizes the relationship between BRICS national S&T thrusts, country scientific specialization and the thematic allocations of the Cape Town Declaration. Subject area publication counts (SC) are displayed as percentages of each country total using the Web of Science ‘Core Collection’ database and Scopus. Also included is the percentage of SC for the fields of astronomy and astrophysics.

Table 1: Thrusts, publication count (2014) and BRICS S&T allocations

SUBJECT COUNT WEB OF SCIENCE	%	BRAZIL	SUBJECT COUNT: SCOPUS	%
Agriculture	7		Medicine	33
Chemistry	7		Agriculture & bioscience	20
Physics	6		Engineering	12
Engineering	6		Biochem, genetics & molecular biology	11
Biochemistry & molecular biology	4		Physics & astronomy	10
THRUSTS :Biotech; Nanotech; Energy; ICT; Health; Biodiversity & Amazon; Climate change; Space science; National security (MST, 2007)				

BRICS S&T theme: Climate change and disaster mitigation				
SUBJECT COUNT: WEB OF SCIENCE	%	RUSSIA	SUBJECT COUNT: SCOPUS	%
Physics	25		Physics & astronomy	32
Chemistry	16		Engineering	20
Material science	7		Material science	18
Engineering	7		Chemistry	14
Mathematics	6		Mathematics	11
THRUSTS: Energy; Nuclear, Strategic ICT; Health; Space science; (Meissner et al, 2013) BRICS S&T theme: Water resources & pollution treatment				
SUBJECT COUNT: WEB OF SCIENCE	%	INDIA	SUBJECT COUNT: SCOPUS	%
Chemistry	19		Medicine	24
Physics	12		Engineering	22
Engineering	11		Computer science	16
Material science	10		Chemistry	14
S&T other	6		Physics & astronomy	13
THRUSTS: Agric, Health, Energy, Transport & infrastructure; Environment; Inclusion; Space (Hoareau McGrath et al, 2014) BRICS S&T theme: Geospatial technologies and applications				
SUBJECT COUNT: WEB OF SCIENCE	%	CHINA	SUBJECT COUNT: SCOPUS	%
Chemistry	18		Engineering	39
Engineering	14		Medicine	18
Material science	12		Material science	16
Physics	11		Physics & astronomy	15
S&T other	8		Computer science	13
NATIONAL THRUSTS: Biotechnology; Food security; Energy sources /materials; Clean vehicles; Climate change/environment (Hoareau McGrath et al, 2014) BRICS S&T theme: New & renewable energy and energy efficiency				
SUBJECT COUNT: WEB OF SCIENCE	%	SOUTH AFRICA	SUBJECT COUNT: SCOPUS	%
Chemistry	6		Medicine	33
Env sci & ecology	6		Agriculture & biosciences	20
Infectious diseases	5		Engineering	12
Engineering	5		Biochem, genetics & molecular biology	11
Physics	5		Physics & astronomy	10
NATIONAL THRUSTS: Biotechnology; Renewable energy; Climate change; Poverty alleviation; Space S&T (DST, 2008) BRICS S&T theme: New & renewable energy and energy efficiency				

Table 2 displays 2014 data from the Web of Science for total publication counts, per BRICS country and the percentage share of the total count in the top four SCs plus astronomy.

Table 2: BRICS research areas and h-indices, Web of Science Core Collection, 2014

BRAZIL 51639		RUSSIA 39963		INDIA 77369		CHINA 329976		SOUTH AFRICA 15337	
Research Area %	h	Research Area %	h	Research Area %	h	Research Area %	h	Research Area %	h
Engineering 8.50	11	Physics 26.00	30	Engineering 17.73	24	Engineering 21.75	41	Engineering 7.88	11
Agriculture 7.25	10	Chemistry 14.21	20	Chemistry 16.50	29	Material sci 16.24	54	Environment 5.89	12
Chemistry 6.99	15	Engineering 10.09	12	Physics 12.32	24	Chemistry 14.80	51	Chemistry 5.38	13
Physics 6.54	22	Material sci 8.50	7	Material sci 10.45	21	Physics 9.99	52	Physics 5.33	17
-		-		-		-		-	
-		-		-		-		-	
Astronomy 1.98	20	Astronomy 4.10	35	Astronomy 1.57	35	Astronomy 0.67	24	Astronomy 3.39	36

The country h-index for each SC is also calculated. This calculation is effected for a particular SC by displaying all country publications in a specific research area ranked according to the number of citations from highest to lowest count. The country h-index is then calculated by inspection according to the standard rule.

Table 3: Leading co-publication count Web of Science (upper diagonal); Scopus (lower diagonal), 2009-2014

	Brazil	Russia	India	China	South Africa
Brazil		<i>Phys 69.0 Astron 27.2</i>	<i>Phys 46.2 Astron 18.8</i>	<i>Phys 48.6 Astron 20.6</i>	<i>Phys 35.8 Astron 17.9</i>
Russia	Phys & Astro 74.6		<i>Phys 64.1 Astron 35.3</i>	<i>Phys 52.2 Astron 20.4</i>	<i>Phys 50.0 Astron 31.1</i>
India	Phys & Astro 46.4	Phys & Astro 72.1		<i>Phys 31.0 Astron 12.8</i>	<i>Phys 19.4 Astron 11.1</i>
China	Phys & Astro 50.9	Phys & Astro 58.4	Phys & Astro 33.5		<i>Phys 32.0 Astron 13.9</i>
South Africa	Phys & Astro 38.1	Phys & Astro 63.2	Phys & Astro 27.7	Phys & Astro 36.5	

Next is the matter of country co-publication. The leading research area(s) for country pair co-publication are shown as percentages of country pair collaboration in Table 3 - Web of Science (upper diagonal, italicized) and Scopus (lower diagonal).

PRELIMINARY RESULTS AND FURTHER ANALYSIS

Kahn (2015) noted a loose correspondence between country national S&T thrusts and BRICS thematic areas as italicized in Table 1. One notes with Yi et al (2013), that with the exception of Russia, the Web of Science categorizations do not indicate overwhelming concentration in one or more fields. Russia's emphasis on physics is the outlier case. On the other hand the Scopus data display much stronger concentrations by SC. This arises from the way that the database architecture combines various sub-fields into larger categories. Russia is once again the outlier in that the leading SC on Scopus, physics and astronomy, 'matches' that found from the Web of Science. There is no obvious match between the thematic allocations of the Cape Town Declaration, and the SC.

Both Brazil and South Africa exhibit a wide spread across research areas, while China, India and Russia display sharper concentration. In all five cases astronomy is an area of low activity.

China has the highest values for the selected h-index. The physics h index is high for all five countries, with the h index for astronomy being as high or much higher in three cases. The main finding is that the BRICS countries would appear to be collaborating in physics and astronomy. While astronomy is the thematic area allocated to South Africa, there is a poor fit between co-publication concentration and the intent of the Cape Town Declaration.

Compared with the five individual BRICS output profiles, whether gleaned from the Web of Science or Scopus, the physics/astronomy dominance is highly concentrated. According to both databases Russia is unique among the BRICS for its revealed specialization in physics.

For Brazil, Web of Science SCs show that physics accounts for 6% of publications, but when it comes to Brazil's co-publication in physics with the other four countries the concentration leaps to 69%, 47%, 48% and 33% respectively. Indeed the median level for physics is 47,5% of all CSCs among the BRICS. For astronomy the BRICS CSC median level is 19,5% of all co-publications.

What then explains this concentration? One possible conjecture is that the concentration arises through the participation of BRICS scientists in the Big Science projects of contemporary physics and astronomy, such as high-energy physics research at CERN, Geneva, and the recent data analysis associated with the Planck satellite observatory through the Planck 2013 Collaboration.

This conjecture is supported by means of the results of a visual scan of the abstracts of publications in physics and astronomy ranked by citation frequency. In physics for example, the publications cited at least 5 or more times are overwhelmingly attributed to large collaborations such as ATLAS, CMS, STAR, ALICE, LHCb etc, while those in astronomy and astrophysics include the Planck 2013 satellite observatory data analysis as well as the high energy physics projects already mentioned. This suggests that the bulk of BRICS collaboration in physics and astronomy takes place via the medium of international Big Science projects.

The conjecture may be tested by examining the publications that contribute to the h index scores for the ten country pairs. The relevant data is provided in Table 4, where the upper diagonal provides the country pair h-index scores, and the lower diagonal the unique number of large

collaborations of the publications that generate the h-index where a named collaboration group is attributed as the author.

Table 4: Co-publication h-index for Physics, Web of Science (upper diagonal); Scopus (lower diagonal) 2014

	Brazil	Russia	India	China	South Africa
Brazil		20	20	20	15
Russia	20		20	17	15
India	20	20		21	15
China	20	17	19		16
South Africa	15	15	15	16	

The concordance between the elements of the upper diagonal and the lower diagonal stands out of the matrix of events.

Table 5: Co-publication h-index for Astronomy and Astrophysics, Web of Science (upper diagonal), Scopus (lower diagonal) 2014

	Brazil	Russia	India	China	South Africa
Brazil		17	16	18	11
Russia	17		35	19	33
India	16	35		16	33
China	18	19	16		13
South Africa	11	33	33	13	

In the case of Astronomy and Astrophysics the data for testing the conjecture are shown as Table 5. One observes that the h-index for the India-Russia, India-South Africa, and South Africa-Russia is much higher than for the other seven country pairs. The explanation is simple: the former three countries are participants in the Planck 2013 Collaboration, whose citations pull up the h-index. In all three cases all but two of the collaborations are through the Planck 2013 analysis. The other seven country pairs include countries that are not involved in Planck 2013 work, and their co-publication citations arise from the inclusion of high-energy physics publications with its overlap into the research area of astrophysics.

The conjecture that collaboration in physics or astronomy/astrophysics is dominated by Big Science projects is essentially confirmed.

DISCUSSION

The BRICS countries have laid out a declaration for collaboration in five specific fields, yet collaboration is currently dominated through the mechanism of pre-existing Big Science projects.

How shall one understand the influence of participation in ‘Big Science’ projects on co-publication, let alone cooperation? Hand (2010) and Ebrahim et al (2013) point to the

(desirable) increase in publication counts resulting from international collaboration, while Birnholtz (2006, 2008) gives specific attention to the problem of attribution in the large projects at CERN, such as the ATLAS. Such issues appear in all large research projects, including astronomy and the health sciences. Hogg et al (2014) point to the emergence of well-crafted protocols that guide article writers in their task. Dance (2012) further discusses the issue of who should be the lead author. None of these contributions consider the distortion of publication counts arising from the massification of research effort as in the 900 student authors¹ of a paper on the genome of the fruit fly.

Participation in particle physics comes at marginal cost, being available remotely. The same holds for Planck satellite data or the Sloan Digital Sky Survey. One can work on ATLAS anywhere, any time, provided that one has the necessary disciplinary knowledge and skills, and are able to get online. The typical ATLAS publication involves hundreds of ‘authors’ most of whom do not know one another, who will never meet, and never correspond. ATLAS Authorship policy is carefully specified.² Indeed it is a condition of working with ATLAS data that all participants are named, so that huge authorship is designed into the governance protocol.³ Moreover the official style guide (Eisenhandler, 2013) requires that the word ‘ATLAS’ must be included in the article title, hence the ease of search.

The question this raises is what significance should be attached to co-publication as a proxy for cooperation? The ‘high’ levels of intra BRICS collaboration in physics and astronomy are in part driven by the protocols pertaining to working with Big Science data.

Should large-scale collaborations be placed in a separate category lest their inclusion acts to distort and exaggerate actual peer-to-peer interaction? Should it be mandatory to introduce fractional counts when an article involves more than a specified minimum number of contributors?

A first step toward resolving these issues and thereby assisting the use of the data in science policy could be to introduce a category ‘Big/Mega Science collaboration’ and to separate off publications that arise from such participation and that have an authorship of ten or more different countries. This category might include both SC and fractional SC, and would be tabulated alongside the SC and fractional SC for all other fields combined.

This is not to suggest that the BRICS countries should limit their involvement in Big Science projects let alone restrict the development of such projects in their own countries. One thinks for example of large-scale nuclear research at Dubna in Russia, the new optical telescope in China, and the incipient Square Kilometre Array in South Africa. What these projects have in common is the generation of Big Data. The capacity of innovation systems large and small, advanced or emergent, in the BRICS or elsewhere, to clean, process and analyse Big Data is now a core attribute for their success into the future. The development and sharing of the appropriate tools arises naturally through international collaboration, even where the collaborators operate remotely attaining ‘satisfaction at a distance.’

¹ <http://www.sciencedaily.com/releases/2015/05/150511095353.htm>

² <https://twiki.cern.ch/twiki/bin/view/Main/ATLASAuthorshipPolicy>

³ The author is grateful to an anonymous physicist with work experience at CERN for these insights.

REFERENCES

- Birnholtz, J. (2006). What does it mean to be an author? The intersection of credit, contribution, and collaboration in science. *Journal of the American Society for Information Science and Technology*, 57(13), 1758–1770.
- Birnholtz, J. (2008). When Authorship Isn't Enough: Lessons from CERN on the Implications of Formal and Informal Credit Attribution Mechanisms in Collaborative Research. *The Journal of Electronic Publishing*, 11(1), Winter 2008. DOI: <http://dx.doi.org/10.3998/3336451.0011.105>
<http://embor.embopress.org/content/embor/10/1/2.full.pdf>
- Bornmann, L., Wagner, C. and Leydesdorff, L. (2015). BRICS countries and scientific excellence: A bibliometric analysis of most frequently-cited papers. *Journal of the Association for Information Science and Technology*, 66 (7):1507–1513. J
- BRICS (2014). BRICS Science, Technology and Innovation Cooperation: A Strategic Partnership for Equitable Growth and Sustainable Development. <http://www.brics.utoronto.ca/docs/140210-BRICS-STI.pdf>. Accessed 2 September 2015.
- Dance, A. (2012). Authorship: Who's on first? *Nature*, 489, 591-593. doi:10.1038/nj7417-591a.
- DST (2008). *Ten-Year Plan for Innovation*. Pretoria: Department of Science and Technology.
- Ebrahim, N. A., Salehi, H., Embi, M. A., Tanha, F. H., Gholizadeh, H., Motahar, S. M., and Ordi, A. (2013). Effective Strategies for Increasing Citation Frequency, *International Education Studies*, 6 (11), 93-99.
- Eisenhandler, E. (2013). ATLAS Style Guide Version 2.4. London: Queen Mary College. <https://cds.cern.ch/record/1110290/files/gen-pub-2008-001.pdf>. Accessed 3 September 2015.
- Finardi, U. (2015). Scientific collaboration between BRICS countries. *Scientometrics*, 102,1139–1166.
- Hand, E. (2010). 'Big science' spurs collaborative trend. Published online 20 January 2010, *Nature*, 463 (282), doi:10.1038/463282a
- Hogg, W., Donskov, M., Russell, G., Liddy, C., Johnston, S. and Mayo-Bruinsma, L. (2014). Approach to publishing for large health services research projects, *Canadian Family Physician*, 60(9), 854–855.
- Kahn, M. J. (2015). Cooperation in science, technology and innovation: rhetoric and realities. *Contexto int.* [online], 37 (1), 185-213. <http://dx.doi.org/10.1590/S0102-85292015000100006>. Accessed 2 September 2015.
- Katz, J. S., and Martin, B.R. (1997). What is research collaboration? *Research Policy*, 26, 1-18.
- Kumar, N. and Asheulova, N. (2011). Comparative analysis of scientific output of BRIC countries. *Annals of Library and Information Studies*, 58, 228-236.
- Meissner, D., Gokhberg, L. and A Sokolov (Eds.) (2013). *Science, Technology and Innovation Policy for the Future*. Berlin: Springer Verlag.
- MST (2007). Science, Technology and Innovation for National Development. Action Plan 2007-2010, Brasilia, Ministry of Science and Technology, 2007.
- Thomson Reuters (2015). <http://wokinfo.com/essays/globalization-of-web-of-science/>

- Tijssen, R., Waltman, L. and Van Eck, N. (2012). Research collaboration and the expanding science grid: Measuring globalization processes worldwide. <http://arxiv.org/abs/1203.4194>
- UNESCO (2010). *World Science Report 2010*. Paris: UNESCO
- Wagner, C. 2008. *The new invisible college: science for development*. Brookings Institution: Washington, DC, USA.
- Wagner, C. and Wong, S. K. (2012). Unseen science? Representation of BRICs in global Science, *Scientometrics*, 90, 1001–1013. DOI 10.1007/s11192-011-0481-z
- Waltman, L., Tijssen, R. and Van Eck, N. (2011) Globalisation of science in kilometres. *Journal of Informetrics*, 5(4), 574-582.
- Yi, Y., Qi, W. and Wu, D. (2013). Are CIVETS the next BRICs? A comparative analysis from scientometrics perspective, *Scientometrics*, 94, 615–628.
- Yang, L. Y., Yue, T., Ding, J. L. and Han, T. (2012). A comparison of disciplinary structure in science between the G7 and the BRIC countries by bibliometric methods. *Scientometrics*, 93, 497–516.



Autonomy vs. dependency of scientific collaboration in scientific performance¹

Zaida Chinchilla-Rodríguez^{*}, Sandra Miguel^{**}, Antonio Perianes-Rodríguez^{***}, María-Antonia Ovalle-Perandones^{***} and Carlos Olmeda-Gómez

^{*} zaida.chinchilla@cchs.csic.es

Instituto de Políticas y Bienes Públicos. Consejo Superior de Investigaciones Científicas, , C/ Albasanz, 26-28. 28037 Madrid (Spain), SCImago Research Group

^{**} smiguel@fahce.unlp.edu.ar

Instituto de Investigaciones en Humanidades y Ciencias Sociales, IdIHCS (UNLP-CONICET), Universidad Nacional de La Plata, calle 51 e/124 y 125, Ensenada, 1925, Argentina, SCImago Research Group

^{***} antonio.perianes@uc3m.es, antonia.ovalle@uc3m.es, carlos.olmeda@uc3m.es

Departamento de Biblioteconomía y Documentación, Universidad Carlos III, , C/ Madrid, 128. 28903 Getafe (Spain), SCImago Research Group

ABSTRACT

This article explores the capacity of Latin America in the generation of scientific knowledge and its visibility at the global level. The novelty of the contribution lies in the decomposition of leadership, plus its combination with the results of performance indicators. We compare the normalized citation of all output against the leading output, as well as scientific excellence (Chinchilla, et al. 2016a; 2016b), technological impact and the trends in collaboration types and normalized citation. The main goal is to determine to what extent the main Latin American producers of scientific output depend on collaboration to heighten research performance in terms of citation; or to the contrary, whether there is enough autonomy and capacity to leverage its competitiveness through the design of research and development agendas. To the best of our knowledge this is the first study adopting this approach at the country level within the field of N&N.

INTRODUCTION

This past decade has staged numerous studies and debates about policy-making processes for science, technology and innovation in Latin America (Campos, Piñero & Figueroa, 2011). Some authors look into the factors conditioning the practice of science in "peripheral contexts", such as selection of research topics, agenda, or criteria behind publication and evaluation (Vessuri, 1984; Kreimer, 2000; 2006). A tension between national and international science is reflected in the research evaluation models, where the criterion based on productivity and publication in peer-reviewed journals of international circulation stands as a bias for the research agendas of the South with respect to the North (Sutz, 2005). López & Taborga (2013) identify international co-authorship, programs of cooperation and scientific mobility as the elements indicating an internationalization of Latin American research. Oregioni (2014) perceives two types of internationalization: direct, established through cooperative relations among researchers, and indirect, promoted by national science organizations by means of evaluation parameters.

¹ This work was made possible through financing by the Project NANOMETRICS (Ref. CSO2014-57770-R) supported by Ministerio de Economía y Competitividad of Spain.

Relationships of scientific co-operation among countries and processes of internationalization have also been approached from the standpoint of academic dependence, understood as an unequal structure of output and divulgation of knowledge on the part of industrialized countries as opposed to peripheral ones (Beigel & Sabea, 2014). Within this framework, the concept of autonomy has been broadly interpreted. Some authors hold that peripheral knowledge is the result of a captive mind; others demonstrate that a peripheral community may reduce its import of foreign knowledge and increase the local production of concepts or methods.

At present, the scientific field of nanoscience and nanotechnology (N&N) has attracted the attention of policymakers worldwide, and several countries have included N&N research programs in their agendas. Its economic and social advances affect sectors such as industry, health, the environment and national security (Huang, et al. 2011). This implies both challenges and opportunities for other countries at medium levels of development, to harness their capabilities and become better situated to benefit from commercial opportunities through targeted investments and strategic collaborations. The rapid growth of N&N is also reflected in the number of publications and patents entailing advancements in knowledge or industrial applications. Against this background, scientific benchmarking can be seen as a useful aid in decision-making about research performance -especially in the case of Latin America- for two main reasons. The first is tied to the "models of academic dependency" outlined above, while the second would be the scarce representation of Latin America in the international arena due to the fact that most studies that analyze this field focus on developed countries (Foladori 2005; Kay & Shapira, 2009; 2011).

This article explores the capacity of Latin America in the generation of scientific knowledge and its visibility at the global level. The novelty of the contribution lies in the decomposition of leadership, plus its combination with the results of performance indicators. We compare the normalized citation of all output against the leading output, as well as scientific excellence (Chinchilla, et al. 2016a; 2016b), technological impact and the trends in collaboration types and normalized citation. The main goal is to determine to what extent the main Latin American producers of scientific output depend on collaboration to heighten research performance in terms of citation; or to the contrary, whether there is enough autonomy and capacity to leverage its competitiveness through the design of research and development agendas. To the best of our knowledge this is the first study adopting this approach at the country level within the field of N&N.

OBJECTIVES AND RESEARCH QUESTIONS

The objective is to characterize the volume, impact, internationalization, scientific capacity and degree of excellence and to evaluate the scientific levels of autonomy or dependency based on internationalization and leadership. We attempt to answer:

- Which countries have greater knowledge output in N&N and a greater degree of leadership and international collaboration?
- Do the levels of scientific performance of a country match its high(er) global output?
- What levels of scientific autonomy and dependency are found when looking at leadership and international collaboration?

MATERIAL AND METHODS

The data set was obtained from SCImago Journal & Country Rank (SCImago 2007) and SCImago Institutions Rankings (SCImago 2013), based on the Scopus database. The indicators used are:

- Output (ndoc): Number of documents published by country.
- Percentage of documents published by Latin American countries in N&N (%LAC)
- Leadership: Percentage of documents published by a country in which the "Corresponding author" is affiliated to a national institution (%lead) (Moya et al. 2013).
- Collaboration types (percentages): a) No-collaboration (non-collab): papers published by one single institution; b) International collaboration: co-authored papers with foreign institutions; and c) International collaboration with leadership: co-authored papers with foreign institutions acting as "Corresponding author"
- Normalized citation impact (NI): The relative number of citations received by each country, compared with the world average of citations received by a paper of the same document type, year and category (Rehn & Kronman 2008; González-Pereira et al. 2010).
- Normalized citation impact with leadership (NIL): this indicator limits its analysis to the leading output
- Benefit rate of collaboration in normalized citation impact (BRCNI), in scientific excellence (BRCE) and in innovative knowledge (BRCIK): the percentage difference between the Normalized Citation/Excellence/Innovative Knowledge of all output and leading outputs. This indicator acts as a proxy to determine the benefit reaped by a country in these indicators when collaboration is not led by the given country. When the value is very low or even negative, it means that the country does not derive much benefit from the collaborations that it does not lead. It signals scientifically well-developed countries whose NI/Exc/IK of total output adequately reflects their scientific performance. A high difference points to scientifically developing countries that depend largely on collaborations with other countries in order to improve their performance. The threshold can vary from one domain to another, but the rule of thumb is: the lesser the benefit rate, the better developed and more autonomous the country.
- Excellence rate (% Exc.): Percentage of documents included within the set of the 10% most cited papers in that category. The percentages can be compared with the "world expected" value established for the top 10% (Tijssen et al. 2002; Bornmann et al. 2012).
- Excellence with leadership (%EwL): Percentage of documents of excellence considered as main contributor.
- Innovative Knowledge: number of all (IK) and leading papers (IK_L) cited in patents.
- Technological impact (%IK): percentage of documents cited in patents with respect to the total output.
- Leadership in technological impact: percentage of leading papers cited in patents with respect to the total output.

RESULTS

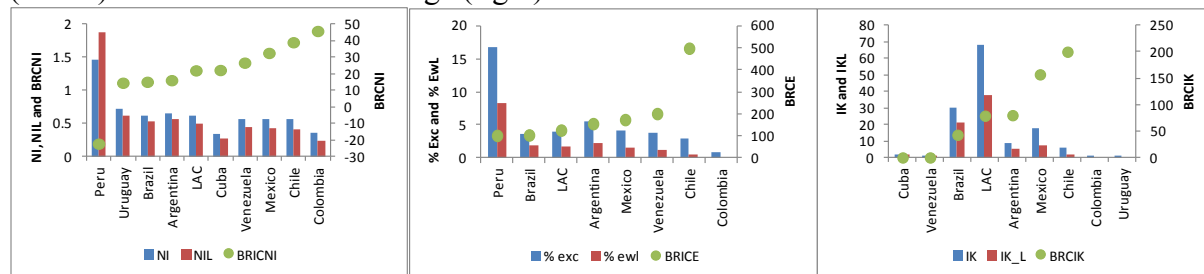
Latin America published 4,811 documents in the category N&N over the time period 2003-2013. This figure represents 2.73% of the world output. Meanwhile, at the world level, a total of 176,158 N&N documents were indexed in Scopus, representing 1.07% of all documents. Brazil is the country with the most output accumulating 46% of the region’s production, followed by Mexico and Argentina (28.46% and 12.51%, respectively). As medium producers, Colombia and Chile contribute similar shares of the regional outputs (roughly 5% and 4%).

Table 1. Main indicators of Latin American countries in Nanoscience and Nanotechnology

Country	ndoc	% LAC	% lead	% non-collab	%lead_collab	% IC	% IC_L	NI	NIL	BRCNI	% exc	% ewl	% eic	BRICE	IK	% IK	IK_L	BRCIK
Brazil	2214	46.02	77.91	28.57	49.34	41.01	24.52	0.61	0.53	15.09	3.66	1.81	1.85	102.21	30	1.4	21	42.9
Mexico	1369	28.46	73.05	32.26	40.79	49.09	30.30	0.57	0.43	32.56	4.16	1.53	2.63	171.90	18	1.3	7	157.1
Argentina	602	12.51	68.77	5.06	63.71	56.98	37.44	0.65	0.56	16.07	5.48	2.16	3.32	153.70	9	1.5	5	80.0
Colombia	235	4.88	68.94	12.01	56.93	67.66	53.09	0.35	0.24	45.83	0.85	0.00	0.85		1	0.4	0	
Chile	210	4.36	60.00	13.11	46.89	70.95	51.59	0.57	0.41	39.02	2.86	0.48	2.38	495.83	6	2.9	2	200.0
Cuba	106	2.20	38.68	12.04	26.64	88.68	70.73	0.33	0.27	22.22					2	1.9	2	0.0
Venezuela	79	1.64	54.43	18.04	36.39	73.42	51.16	0.57	0.45	26.67	3.80	1.27	2.53	199.21	1	1.3	1	0.0
Uruguay	38	0.79	60.53	46.24	14.29	57.89	30.43	0.71	0.62	14.52						0.0		
Peru	12	0.25	16.67	0.00	16.67	100.00	100.00	1.46	1.88	-22.34	16.67	8.33	8.34	100.12	1	8.3		
LAC	4811	100.00	76.20	23.19	53.01	46.21	30.39	0.61	0.50	22.00	3.97	1.77	2.20	124.29	68	1.4	38	78.9

A close look at the relationship between leadership and collaboration (Table 1) reveals diverse patterns, with certain countries showing a high percentage of documents in total and leading international collaboration, such as Cuba. It is followed by Venezuela, Chile and Colombia, with proportions of CI from 67.7% to 73.4%, and output with leadership in CI over 50%. In other countries, the leadership is concentrated in “inbreeding” production, whereby nearly 70% of the output involves documents of national collaboration or non-collaborative production. The most noteworthy cases are Brazil and Mexico, which have the least international collaboration (below the Latin American average of 30.4%). They are followed by Argentina, whose leadership model is more strongly based on national collaboration: it shows the lowest level of non-collaborative output (5.06%), and a moderate degree of international leadership (37.4%), higher than Brazil and Mexico. In Uruguay, leadership essentially entails output not involving collaboration.

Figure 1. Benefit rate of collaboration in normalized citation (left), scientific excellence (center) and innovative knowledge (right)



Peru presents an interesting profile regarding the benefits of collaboration on the impact and excellence. As shown in Fig. 1, Peru’s output is totally dependent on IC. No other Latin American country reaches impact levels near the world average. Yet the levels of impact attained with leadership (NIL) are lower than the global impact (NI). Similarly, regarding excellence, there are greater differences between the output of excellence and the output of excellence that is led by each country.

Colombia derives the greatest benefit from collaboration in attaining impact. At the other end of the scale, Argentina, Brazil and Uruguay have values below the average, benefit less from collaboration with regard to impact, and are more autonomous in achieving scientific impact, although it remains below the worldwide average (Fig. 1-left). Chile gains the most from collaboration with excellence, and Brazil and Argentina are the ones with the least incidence of collaboration in this segment of output.

As can be seen in Fig. 2 (right), Chile is clearly the country that derives the greatest benefits from collaboration in terms of technological impact, followed by Mexico. Argentina and Brazil are less dependent upon collaboration. Figure 2 reflects the position from the perspective of scientific and technological dependence upon collaboration in order to reach their levels of impact. Again, this signals that Chile and Mexico are hardly independent, calling for collaboration in order to attain scientific or technological impact.

Figure 2. Autonomy vs dependency of collaboration in scientific and technological impact

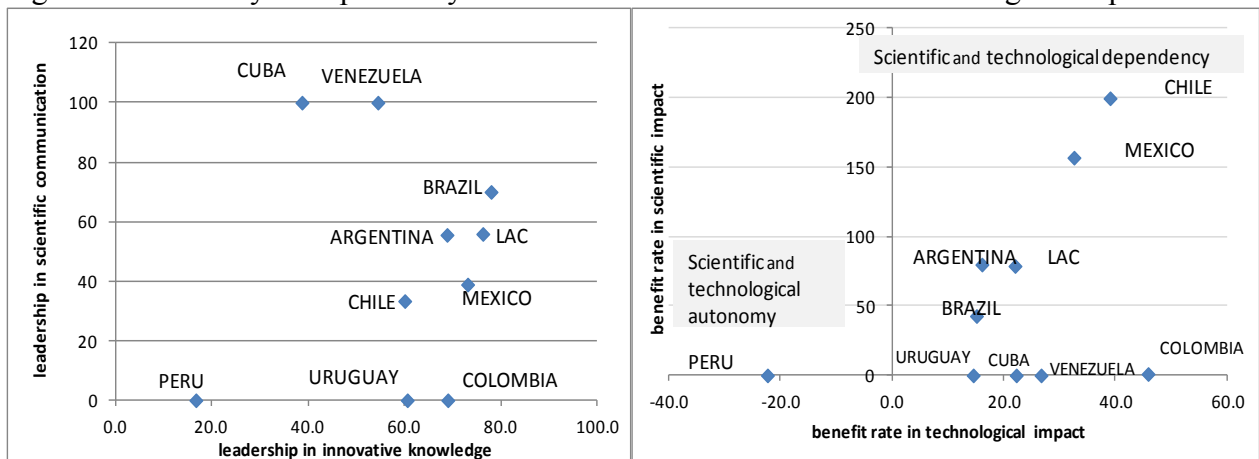
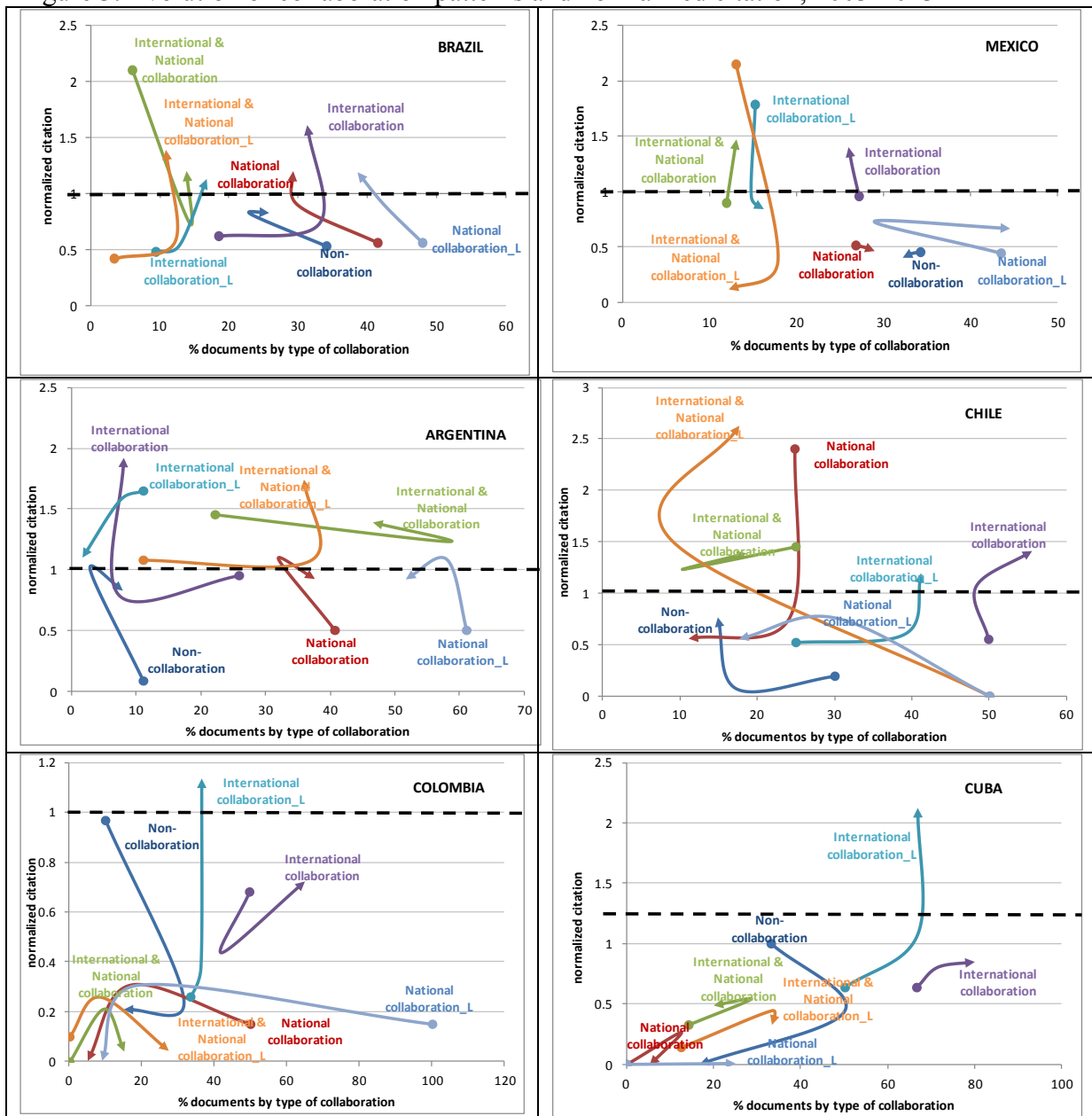


Figure 3 illustrates the trends behind the influence of collaboration in the normalized citation of each country. Brazil presents a clear growing trend in its total and leading international collaboration, as well as in national collaboration and international collaboration with leadership. In all three cases the value has a positive effect on impact, yet international collaboration is the most determinant one for high citation. Furthermore, output in collaboration allows Brazil to eventually obtain normalized citation indexes above the world average. Non-collaboration decreases and has the least impact.

In Mexico the panorama is quite different. There is a rising trend for the impact of output entailing international collaboration, as well as national and international collaboration, whereas a reverse trend is seen for the levels of impact of output coming from leading collaboration, and non-collaboration.

Argentina tends to increase the impact of its output resulting from international collaboration, while the national and international production decline in impact. The other types of output show impact on the rise, by the end of the period over the world average in all cases except non-collaborative output, which holds an impact under 1 throughout the period.

Figure 3. Evolution of collaboration patterns and normalized citation, 2003-2013



Chile's output with international collaboration and in IC-L increase steadily, in both cases reaching impact values above the world mean by the end of the period analyzed. However, it is the output combining national and international leadership and collaboration which achieves the greatest levels of impact, and finally reaches a value over 2.5, despite a slight drop in the % of documents.

The case of Colombia is interesting indeed. All its output except international with leadership has impact levels under the world mean at the end of the period, showing virtually no percentage-wise growth. Thus, this country obtained more benefits from international collaboration with leadership, but did not increase its share of production. It shows no real decline in non-collaborative output, though its impact is substantially reduced. The greatest change is in national and international output, with or without leadership, although in either case the impact achieved lies below the world average.

Cuba only increases its output in international collaboration, and much more in output with leadership than without leadership. This means it also considerably increases the levels of impact, which at the end of the period reach an index over 2. Non-collaboration shows a considerable decline in both output and impact. The national and international collaboration without leadership grows slightly in volume and impact, while collaboration with leadership grows in volume but ends up with somewhat less impact.

CONCLUSIONS

Even if the argument, some theoretical assumptions and the consequences of the method employed should be discussed, the approach taken here, with its emphasis on the autonomy and dependency of countries in their performance, tries to contribute to constructive debate about how best to assess size and performance in future studies. This new methodology can be extrapolated to different fields of study. The metrics provided here are not the only indicators that might account for analyzing research performance. It follows that any measure used to gauge impact will also reflect social factors beyond the conventions or patterns of behavior of scientific output. As strongly advocated in the Leiden manifesto, scholarly metrics should play the supporting role to qualitative and in-depth analyses of scholarly content and activities (Hicks et al. 2015). There are many ways to expand upon this analysis in order to enrich and complement the findings exposed here.

Taking into account that leadership means responsibility and acknowledgment of the responsible for the publication of research, leadership also means merit when it entails international collaboration. Thus, leadership and international collaboration patterns help to characterize how research is carried out, taking into account scientific capacities in linkage networks and to what extent countries play different roles in the management of their own capacities to generate knowledge and to attract international partners (Chinchilla et al, 2016a; 2016b).

The growth of international collaboration with leadership should be interpreted as a positive aspect, a progressive internationalization of scientific activity with capacities for the definition of research agendas concerned with local needs or topics of interest that would likewise be of interest for further communities abroad, in turn contributing to the development of science at a national level. Although we need to explore scientific and technological leadership in greater depth, this statement attempts to sum up the significance of so-called tensions between the autonomy/dependence of scientific agendas and how they might be remedied by greater opportunities to increase the visibility and competitiveness of research in peripheral countries. Therefore, the data stand as an invitation for researchers to carry out studies in greater depth and identify the groups and subject areas in which Latin America demonstrates greater potential.

REFERENCES

- Beigel, F., Sabea, H. (coords.) (2014). *Dependencia académica y profesionalización en el sur. Perspectivas desde la periferia*. Mendoza: EDIUNC; Rio de Janeiro: SEPHIS.
- Bornmann, L., Moya-Anegón, F. & Leydesdorff, L. (2012). The New Excellence Indicator in the World Report of the SCImago Institutions Rankings 2011. *Journal of Informetrics*, 6, 333–335.
- Campos, G., Piñero, F. & Figueroa, S.A. (coords.) (2011). *Transformaciones recientes de las universidades latinoamericanas. Agendas y actores en la producción de conocimiento*. México; Tandil: Benemérita Universidad Autónoma de Puebla; Universidad Nacional del Centro de la Provincia de Buenos Aires.
- Chinchilla-Rodríguez, Z., Zacca-González, G., Vargas-Quesada, B. & Moya-Anegón, F. (2016a). Benchmarking Scientific Performance by Decomposing Leadership of Cuban and Latin American Institutions in Public Health. *Scientometrics*, 106, 1239-1264.
- Chinchilla-Rodríguez, Z., Ocaña-Rosa, K., & Vargas-Quesada, B. (2016b). How to combine research guarantor and collaboration patterns to measure scientific performance of countries in scientific fields: nanoscience and nanotechnology as case study. *Frontiers in Research Metrics and Analytics*, doi: 10.3389/frma.2016.00002
- Foladori, G. (2006). Nanotechnology in Latin America at the crossroads. *Nanotechnology Law & Business*, 3, 205–216
- González-Pereira, B., Guerrero-Bote, V. P. & Moya-Anegón, F. (2010). A New Approach to the Metric of Journals' Scientific Prestige: The SJR Indicator. *Journal of Informetrics*, 4, 379–391.
- Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. (2015). Bibliometrics: the Leiden Manifesto for research metrics. *Nature*, 520, 429-431.
- Huang, C., Notten, A., & Rasters, N. (2011) Nanoscience and technology publications and patents: a review of social science studies and research strategies. *The Journal of Technology Transfer*, 36, 145-172
- Kay, L., & Shapira, P. (2009). Developing Nanotechnology in Latin America. *Journal of Nanoparticle Research*, 11, 259–278.
- Kay, L., & Shapira, P. (2011). The Potential of Nanotechnology for Equitable Economic Development: The Case of Brazil. *Nanotechnology and the Challenges of Equity, Equality and Development*. Springer, 309–329.
- Kreimer, P. (2000). Cultura y periferia. La ciencia en la Argentina entre siglos. En Montserrat, Marcelo, *Textos, contextos e instituciones*. Buenos Aires: Manantial, 187-202.
- Kreimer, P. (2006). ¿Dependientes o integrados? La ciencia latinoamericana y la nueva división internacional del trabajo. *Nómadas*, 24, 190-212.

- López, M.P. & Taborga, A.M. (2013). Dimensiones internacionales de la ciencia y la tecnología en América Latina. *Latinoamerica*, 56, 27-48.
- Moya-Anegón, F., Guerrero-Bote, V.P., Bornmann, L. & Moed, H.F. (2013). The Research Guarantors of Scientific Papers and the Output Counting: A Promising New Approach. *Scientometrics*, 97, 421–34.
- Oregioni, M.S. (2014). Incidencia directa e indirecta de la internacionalización de la investigación en la Universidad Nacional de La Plata, desde una perspectiva multidimensional. *VIII Jornadas de Sociología de la UNLP*, 2014.
- Rehn, C., & Kronman, U. (2008). *Bibliometric Handbook for Karolinska Institutet*. Huddinge:Karolinska Institutet.
- SCImago. (2007). *SCImago Journal & Country Rank*. <http://www.scimagojr.com>.
- SCImago. (2013). *SCImago Institutions Rankings*. <http://www.scimagoir.com>.
- Sutz, J. (2005). Sobre agendas de investigación y universidades de desarrollo. *Revista de Estudios Sociales*, 22, 107-115.
- Tijssen, R. J. W., Visser, M. S., & van Leeuwen, T. N. (2002). Benchmarking International Scientific Excellence: Are Highly Cited Research Papers an Appropriate Frame of Reference? *Scientometrics*, 54, 381–397.
- Vessuri, H.M.C. (ed.) (1984). *Ciencia académica en la Venezuela moderna*. Caracas: Fondo Editorial Acta Científica Venezolana.



The world network of scientific collaborations between cities: domestic or international dynamics?¹

Marion Maisonobe*, Denis Eckert*, Michel Grossetti**, Laurent Jégou*, Béatrice Milard**

*marion.maisonobe@univ-tlse2.fr, eckert@univ-tlse2.fr, jegou@univ-tlse2.fr

UMR LISST-CIEU, CNRS, Université de Toulouse, 5 allées Antonio Machado, Toulouse, 31058 (France)

**rgros@univ-tlse2.fr; milard@univ-tlse2.fr

UMR LISST-CERS, CNRS, Université de Toulouse, 5 allées Antonio Machado, Toulouse, 31058 (France)

ABSTRACT

Earlier publication (Grossetti et al., 2014) has established that we are attending a decreasing concentration of scientific activities within “world-cities”. Given that more and more cities and countries are contributing to the world production of knowledge, this article analyzes the evolution of the world network of collaborations both at the domestic and international levels during the 2000s. Using data from the Science Citation Index Expanded, scientific authors’ addresses are geo-localized and grouped by urban areas. Our data suggests that interurban collaborations within countries have increased together with international linkages. In most countries, domestic collaborations have increased faster than international collaborations. Even among the top collaborating cities, sometimes referred to as “world cities”, the share of domestic collaborations is gaining momentum. Our results suggest that, contrary to common beliefs about the globalization process, national systems of research have been strengthening during the 2000s.

INTRODUCTION

The global “growth of science”, world-wide access to transportation, information, and communication technologies, as well as collaborative research policies, have encouraged international scientific cooperation. Together with the continuing spatial diffusion of scientific activities at the world level (Inhaber, 1977; Grossetti *et al.*, 2014), the increase of scientific collaboration is often described as one of the main features of globalization (Schott, 1993; Wagner, 2008; Royal Society, 2011; Sexton 2012). Measuring the growth of scientific collaboration both within and across countries during the 2000s, and taking into account the share of intercity co-authorships, this article provides new evidence regarding the evolution of the world collaboration network.

Previous works used to focus only on the top publishing or cited urban areas in the world (Matthiessen *et al.* 2010; Bornmann *et al.*, 2011). They also used to limit their scope to certain macro regions such as Europe (Zitt *et al.*, 1999; Hoekman *et al.*, 2009). Our contribution is a spatial analysis of the co-authored articles, reviews, and letters extracted from the *Science Citation Index Expanded* with the much sharper spatial resolution of the urban level. Here, we

¹This work was supported by the French Research Agency (L’Agence Nationale de la Recherche) through the GEOSCIENCE project and the Excellence laboratory “Structuring of social worlds” (ANR-11-LABX-0066) funded by the French Research Agency (Agence Nationale de la Recherche) as part of the program n° ANR-11-IDEX-0002-02. A longer version of this article will soon be published in the journal *Informetrics*.

want to answer the question: what are the territorial dynamics underlying the growth of interurban collaborations during the last decade?

First, we present our spatial bibliometrics method based on the geocoding of publications and their assignment to urban areas. Then, we show that the more developing is an area (city and country), the more scientists located in this area have favored domestic collaborations. To finish, we demonstrate that national systems have remained highly structuring, even when considering only the position of the top world cities.

2. MATERIALS AND METHODS

In order to measure the scientific activities of various cities, the more robust approach is to process the institutional affiliations (addresses) contained in bibliometric data. Since the 2010s, much progress has been made in processing the spatial information of bibliometric data at a higher level of resolution (Leydesdorff & Persson, 2010). Performing a bibliometric analysis of urban production has been worthwhile to find that the world scientific production is performed by an increasing number of cities (Grossetti *et al.*, 2014). What is happening is that the previous monopoly of capital cities or historical university/research centers is, little by little, diminishing in almost every country in the world. Exploiting also geolocalization tools, Tijssen *et al.* have proven that the mean kilometric collaboration distance has increased globally during the 2000s whereas the share of international collaboration(s) has leveled off (Tijssen *et al.*, 2012). Drawing upon this last family of collaboration studies, the multi-level analysis approach we propose can be used to describe the evolution of the world collaboration network at an unprecedented level of geographical resolution: the urban area level (urban agglomerations, that is to say perimeters merging cities with their suburban areas). Using urban area perimeters that we designed to this aim, we can distinguish globally between two kinds of interurban collaborations, domestic collaborations and international collaborations and wonder whether or not the growth of scientific collaborations between cities has been influenced by national dynamics and politics.

2.1. The geocoding process and the building of scientific agglomerations

The quality of automatic geocoding tools (Google, Yahoo!, Bing, etc.) is actually widely divergent when used on a worldwide and spread over several decades dataset, such as the one we have used: the *Science Citation Index Expanded*. In this bibliometric database, the authors' addresses are decomposed in several fields, of which we selected three: locality, province, and country. Our target scale for geocoding was the locality level. Error control and correction was quite a long procedure, helped by the development of a user-friendly online visualization tool shared among all project participants. For instance, we detected an erroneous location in Southern Germany. The “Garching” text string, which refers to a suburb of Munich harboring a huge science & technological park, was first located by the geocoding tool in remote rural Bavarian area. A comparable problem was visually detected near Chicago (Argonne). A data-quality index was constructed by country, indicating the zones where expert verification was needed. The quality of the geocoding improved step by step. After more than a year of work, with the help of geospatial analysts and cartographers working in fields such as sociology and geography of science, we obtained a fine-tuned/high-resolution spatial database of scientific production over the last decade.

This granularity is itself a source of problems when attempting a comparative approach at the global level. The characteristics of postal addresses, the geographical variability of postal reference systems, and the great diversity of administrative geographical segmentation prevent any direct comparison between the 18 650 distinct “scientific localities”. Our group addressed this problem by merging localities into urban areas. The goal was to build spatially

comparable geographical entities at the global level. Once all articles, reviews and letters were extracted from the *Science Citation Index* from two time series (1999-2001; 2006-2008) and geocoded, 10 730 urban perimeters were delineated and used as elementary analysis units to measure scientific activity.

The method we used to build those entities is described in more details in Grossetti *et al.* (2014, *op. cit.*). For the purposes of this article, it is important to remember that we used a two-step method. First, the city perimeters defined around the 500 top publishing localities were obtained using a supervised procedure based on population density (highly fine-tuned raster data) and scientific production volumes. Second, smaller publishing localities, which were not included within one of the dense urban areas, were merged together if they were geographically close enough, with 40 kilometers as the criterion. By the means of our method, the Parisian agglomeration includes suburban localities such as Gif-sur-Yvette, Villejuif, l'Université de Versailles-Saint-Quentin-en-Yvelines. Once the urban areas were delineated taking into account the localization of scientific activity, a final step was required: the selection of a counting method.

2.2. The counting method

To measure scientific activities it is first necessary to select a counting unit and a counting system. Second, it is advised to adopt a smoothing method based on averages in order to level short term fluctuations, in particular spikes related to the variable periodicity of issues in scientific journals. The method to be chosen varies according to the stakes of the research, the analytical scale, the dataset. Here, our dataset is one of the most reliable to obtain a global overview of the world science. The *Science Citation Index Expanded* (SCIE) indexes international journals from at least 8 scientific fields (engineering, physics, biology, mathematics, chemistry, medicine, science of the universe and multidisciplinary sciences). However, its coverage is biased toward Anglo-American journals as well as biomedical literature.

The counting unit we considered is the urban area instead of the address. This choice means that we are performing a “whole” instead of a “complete” count (Gauffriau *et al.*, 2008). As a result, the participation of cities that we measure is not influenced by the precise number of addresses per city per publication which is a country-dependent factor (Eckert *et al.*, 2013). Doing so, we are simplifying the intra-local information to focus on the interurban activities. This counting method is more robust at the world level since it is not influenced by the worldwide heterogeneity of administrative fragmentation. In order to study collaboration networks and avoid double counts, co-authored publications are normalized using fractional counts. Fractional counting is the best way to avoid double-counts that would disrupt a multiscalar analysis. When studying a co-authorship network, the whole-normalized counting method evaluates each publication according to the number of urban areas it comes from. For instance, if a publication is signed by scientists located in Paris, Villejuif, Toulouse, and London; only 3 urban areas are counted: Paris, Toulouse and London, since Villejuif is part of the Paris agglomeration. The collaboration volume between all pairs of cities will account for 1/3. In this instance, the share of domestic collaboration is 1/3, and the international share accounts for 2/3. By means of this fractioning method, it is possible to create any sum without losing the real number of co-authored publications issued at the world level. Following most studies, we use co-authorship data as a proxy for collaboration data. We consider that an article co-signed by authors located in two different urban areas or signed by an author affiliated to two different urban areas is the trace of an institutional link between both places. This link may have been prompted by the circulation of an idea, of a material or even of the scientist who wrote the article.

3. RESULTS AND DISCUSSION

3.1. The global growth of multi-city publications

During the last decade, how many publications came from a single city? How many came from several cities in the same country? How many came from several cities in several countries? In Table 1 designed to answer those questions, publications can only belong to one category at a time. When a publication is co-signed by groups from several cities, should one of the cities be located in a different country than the others, the publication is registered in the last category: “several cities, several countries”.

At the world level, Table 1 shows a very unambiguous global trend: single-city publications decreased between 2000 and 2007 to the benefit of multi-city publications. The 4 point loss in the share of single-city publications benefited all other types of publications relatively equally: publications from several cities in the same country, and from several cities in different countries. This latest growth was even slightly more pronounced within countries than between them. Nevertheless, the share of intra-urban publications was still larger than that of interurban publications in 2007, which is in line with the scientific literature dealing with proximity effects.

Table 1. The growth of interurban collaborations between 2000 and 2007. Source: SCI Expanded (articles, reviews, letters), 3-year moving average.

Share of scientific publications signed from:	2000*	2007*
One city	68,8	64,9
Several cities in the same country	15,3	17,3
Several cities, several countries	16,0	17,8
Total (%)	100,0	100,0
<i>Total number of articles, reviews, letters</i>	<i>763203,0</i>	<i>1117566,3</i>

Focusing on country-to-country differences in the next section demonstrates that even though all countries saw their share of single-city publications decline, the level of benefit for domestic and international collaborations varied substantially from one country to another. In order to go further in the analysis of interurban collaborations, we perform an analysis based on normalized data (fractional counts).

3.2. Variations in the evolution of international and domestic interurban collaborations

The following results are obtained by focusing on interurban collaborations only, that is to say on 35% of the total production in 2007 (386 255 publications are the result of interurban collaborations in 2007). In Table 2, the number of interurban collaborations is split between domestic and international. In most countries, the domestic share of interurban collaborations increased faster than the international share. The reinforcement of intra-national collaborations is most obvious in emerging countries, more exactly countries which have been through major reconfigurations and upheavals during the last several decades (China, Brazil, India, Taiwan, Poland, Turkey, Greece, Czech Republic, Iran). Our hypothesis is that the reinforcement of national systems of research within these countries has led to more scientific autonomy for each. In other words, scientists in these countries depend less and less on international groups to publish in highly visible journals. Indeed, with the deconcentration

process studied by Grossetti *et al.* (2014), opportunities for intra national collaborations have been opened up.

The few countries where a trend toward the internationalization of scientific collaborations can be clearly evidenced are Anglo-American (Japan and Sweden excepted): United-States, United Kingdom, New Zealand, and South Africa. Since English is the scientists' language in those countries, this trend could result from the transition toward of a trans-national model of communication (more and more authors willing to publish with English-speaking scientists in scientific journals indexed in the SCI during the 1990s).

Table 2. Evolution of the domestic and international share of scientific collaborations at the country level

35 most co-authoring countries (2007*)	% of the country in the whole interurban co-authorship (2007*)	Domestic co-authorship (2000-2007)**		International co-authorship (2000-2007)**	
Czech-Republic	0.5	1.45	↗↗	0.85	↘↘
Poland	1.0	1.38	↗↗	0.85	↘↘
Portugal	0.6	1.36	↗↗	0.85	↘↘
Greece	0.6	1.35	↗↗	0.83	↘↘
Hungary	0.4	1.33	↗↗	0.90	↘↘
Turkey	1.0	1.30	↗↗	0.63	↘↘
Iran	0.5	1.28	↗↗	0.73	↘↘
Brazil	2.0	1.25	↗↗	0.69	↘↘
Russia	1.4	1.23	↗↗	0.90	↘↘
Mexico	0.6	1.20	↗↗	0.89	↘↘
Israel	0.7	1.18	↗↗	0.92	↘↘
Belgium	1.1	1.14	↗↗	0.94	↘↘
Taiwan	1.5	1.13	↗↗	0.74	↘↘
India	1.7	1.12	↗↗	0.84	↘↘
China	7.3	1.10	↗↗	0.82	↘↘
Spain	2.6	1.08	↗	0.94	↘
Argentina	0.4	1.07	↗	0.96	↘
Denmark	0.8	1.07	↗	0.97	↘
Austria	0.7	1.06	↗	0.99	↘
Netherlands	2.1	1.04	↗	0.97	↘
Australia	2.0	1.02	↗	0.99	↘
South-Korea	2.1	1.02	↗	0.97	↘
Finland	0.8	1.02	↗	0.98	↘
Switzerland	1.4	1.01	-	1.00	-
Norway	0.6	1.01	-	0.99	-
Germany	6.8	1.01	-	0.99	-
Italy	4.0	1.01	-	0.99	-
Canada	3.3	1.00	-	1.00	-
France	5.0	1.00	-	1.00	-

United-States	26.7	0.99	↘	1.03	↗
Japan	6.5	0.98	↘	1.06	↗
Sweden	1.4	0.97	↘	1.02	↗
United-Kingdom	6.1	0.90	↘	1.10	↗↗
South-Africa	0.4	0.89	↘↘	1.07	↗↗
New-Zealand	0.4	0.86	↘↘	1.08	↗↗
Total of the 35 countries	95.0	1.03	↗	0.97	↘
World total	100	1.02	↗	0.97	↘

*Normalized counting (WNC), 3-years moving average. Source: SCIEp (articles, reviews, letters)

** % co-authorship 2007 / % co-authorship 2000

Key: The % of domestic interurban co-authorship of Czech-Republic have been multiplied by 1.45 between 2000 and 2007

Similar results are found if we focus on the top publishing cities in the world (Table 3). Few of these cities developed more international links between 2000 and 2007 whereas the majority saw an increase in their domestic share of collaborations. It appears that all “elite” cities follow their national trend (p -value ≤ 0.001); which suggests that even if a city belongs to the ‘rich club’ (according to scientific production), it does not mean the city is impervious to national logics (Table 3). ‘World’ cities characterized by an unambiguous trend toward nationalization can be found in countries where the scientific production has been deconcentrating the most over the last decade, in particular: China, South Korea, Spain and Russia.

Table 3. Evolution of the national share of scientific collaborations at the city level

The top publishing cities in 2007*	Dynamic of the domestic co-authorship (2000-2007)*		Dynamic of the international co-authorship (2000-2007)*	
Moscow	1.27	↗↗	0.92	↘
Taipei	1.21	↗↗	0.76	↘
Hong Kong	1.20	↗↗	0.81	↘
Beijing	1.14	↗↗	0.81	↘
Melbourne	1.11	↗↗	0.94	↘
Madrid	1.11	↗↗	0.94	↘
Munich	1.07	↗	0.95	↘
Barcelona	1.07	↗	0.97	↘
Shanghai	1.05	↗	0.93	↘
Montreal	1.04	↗	0.98	↘
Roma	1.03	↗	0.97	↘
Toronto	1.02	↗	0.99	↘
Seoul	1.01	-	0.99	-
Paris	1.01	-	1.00	-
San Francisco Bay	1.00	-	1.01	-
Berlin	0.99	-	1.00	-
Washington-Bethesda	0.99	-	1.01	-
Durham Research Triangle	0.99	-	1.02	-
Philadelphia	0.99	-	1.02	-
Kyoto-Osaka	0.99	-	1.02	-

Chicago	0.99	-	1.02	-
Milan-Pavia	0.99	-	1.01	-
New York	0.98	↘	1.03	↗
Tokyo	0.98	↘	1.04	↗
Sydney	0.98	↘	1.01	↗
Los Angeles	0.98	↘	1.04	↗
Baltimore	0.98	↘	1.06	↗
Boston	0.95	↘	1.08	↗
London	0.86	↘↘	1.09	↗↗

*Fractional counting (WNC), 3-years moving average. Source: SCIEXP (articles, reviews, letters)
Key: The % of domestic interurban co-authorship of Moscow have been multiplied by 1.27 between 2000 and 2007

To summarize, in most countries, and even among the top publishing cities, the growth of domestic interurban collaborations has exceeded the growth of international linkages. In other words, there has not been any sizable and unilateral trend toward internationalization at the global level during the last decade. We expect those observations to have structural effects on the world collaboration network.

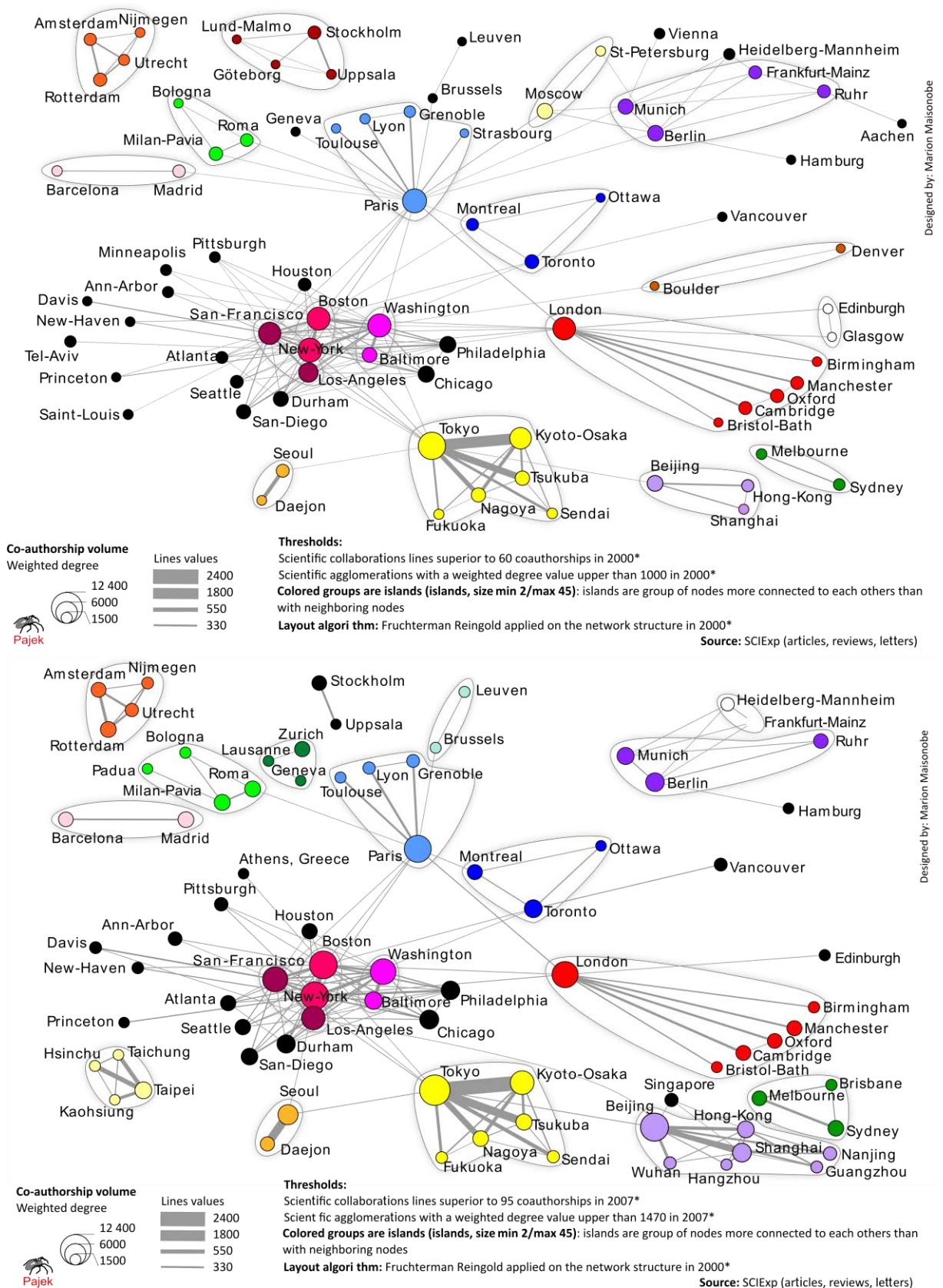
3.3. The evolution of the world collaboration network: densification, deconcentration and strength of domestic links

In order to assess changes in the structure of the world collaboration network, we focus on the co-authorship matrix of the 500 top publishing cities in 2007². These cities are from 59 different countries and are responsible for 87% of world publications in 2007 (960 880 publications). In analyzing the same set of cities in 2000 and 2007, we can track the evolution of the collaboration network not being influenced by the entrance of new production centers during the period. By this method, it appears that the isolated or weakly connected cities in 2000 are much more integrated in 2007.

The global connectivity of the network (the density and the degree indicators) increased by 25% between 2000 and 2007. In 2007, almost every city is related to the others by at least one co-authorship link. As a result, the participation share of the top collaborating cities has decreased within the total collaboration volume. The 100 top collaborating cities were still involved in the majority of collaborations in 2007, but their contribution went from 59% to 57% of all collaborations. Thus, there has been a clear, though slow, trend toward the deconcentration of collaborations. By focusing on the more central cities of the network, we want to examine the impact of this deconcentration process on their connectivity.

² This matrix has been computed on the basis of collaborative articles, reviews, and letters indexed in the SCI Exp in 1999-2001 and 2006-2008. We have chosen to stop at 500 cities because these cities figured as major publishing places in both the entire Web of Science catalogue and the SCI Expanded.

Figure 1 and 2. World network of interurban scientific collaborations in 2000 and 2007 (Whole Normalized Counting, 3- years moving average)



In focusing on the most collaborative cities and on the strongest links between them, we obtain Figure 1 and 2. In each figure, the visible cities are those which, ranked in descending order, are involved in 55% of all interurban collaborations, and the visible links are those which, ranked in descending order, account for 20% of all interurban collaborations. Using a similar criterion for both figures ensures their comparability. The fact that an almost identical number of cities (about 80) is obtained in both figures suggests stability³ but the visible cities are not the same in 2000 and 2007. The scientific collaboration network is first of all expanding in favor of previously peripheral countries (hence the entrance of Chinese cities in the top of the hierarchy).

Interestingly, the figures show that the trend toward nationalization previously identified in some countries is still obvious when focusing on the top interurban collaborations. Indeed, the strongest links in 2007 are more often intra-national than in 2000. Thus, whereas there were four strong components (distinct groups of cities) in 2000 (the main one + Netherlands, Sweden and Australia), there are six of them in 2007 (the previous ones + Switzerland, Spain and Taiwan). Overall, we notice that national systems are reinforcing.

The fact that only a few “world” cities share very strong international links is in line with many studies on world cities’ networks (Matthiessen, 2010). However, contrary to most authors in this field of study, we oppose the idea of a trend toward an increasing concentration of scientific activities within world cities at the expense of smaller cities. Our results suggest that scientific collaborations between cities are strengthening at different levels: regional, national, global.

The network analysis confirms that the growing integration of national systems should not be ignored to understand the geography of scientific activity at the city level.

However, national systems are not the only organizations with structural effects on scientific collaboration networks. On figure 1 and 2, the color of the nodes depends on the “islands” they belong to according to their collaboration patterns. An “island” is a group of cities that share relationships whose values are higher than the strongest bond they have developed outside their group. In order to detect these groups, we have used a clustering method implemented in the Pajek software (used and defined by Batagelj *et al.* (2006) to analyze patent data in the United States). Although in most cases, islands are entire countries, there are certain instances where islands are sub-national groups. Actually, sub-national groups have been detected in the United-States, Germany and the United-Kingdom. Not surprisingly, these countries have a more federal organization than the others so that the sub-national islands we found within them group together cities belonging to the same province (e.g. California and Scotland).

While the growth of interurban scientific collaborations occurred mainly within national frames between 2000 and 2007, the world-wide network expanded globally and locally at the same time.

CONCLUSION

The data analysis of scientific publications between 2000 and 2007, focusing on the development of interurban collaborations, highlights several global trends. The majority of publications were still produced by a single city; but this proportion declined everywhere in the world. There was an overall increase in the number of publications produced by several

³*nota bene*: Large research centers such as Helsinki, Copenhagen or Oslo are not represented since they only match one of the two criteria we used for the visualization. Indeed, they are well connected cities (a high number of collaborations) but they don’t share any collaboration link superior to the selected threshold

cities, within a single country or several. Interestingly, cities located in scientific emerging countries tended to favor domestic interurban co-authorships whereas cities located in more traditionally English-speaking and developed countries did internationalized. Actually, in most countries, there has been a general increase of all kinds of collaborations. As a result, the global interurban network of collaborations densified from 2000 to 2007. Among the top cities, intra-national links reinforced and Asian cities superseded certain Eastern and Northern European cities. All of these results suggest that the worldwide growth of scientific collaborations, referred to as “globalization”, did not develop at the expense of national systems of science. According to us, these results are the consequence of the decentralization process which took place in almost every country following the devolution of higher education services. This devolution process has favored the growth of research by hiring scientists in new scientific centers. Little by little, their collaborative methods tend to become identical to that of the top cities’ scientists. As we saw, the national level has been an essential component for the development and integration of scientific cities. This level has remained equally structuring for more traditional territories where national and international collaborations have complemented each other. Our results highlight the role of national systems of cities in collaboration dynamics during the last decade.

This contribution is part of an ongoing research process. Some limitations should be discussed since they can open up lines for future research. First, our study takes into account all publications without considering the number of citations received. It will certainly be necessary to repeat the analysis by integrating this variable to see if the trends are the same if we examine the most cited publications. Second, our data end in 2008 and it will be necessary to continue the analysis with more recent data, which means to repeat the geocoding and checking the stability of scientific agglomerations. By this means, it will be possible to test the hypothesis according to which collaborative practices in traditional and new centers of scientific activity have continued to converge since 2007, both at the country level and at the urban area level. Third, our network analysis focuses on the question of the hierarchy of the centrality and overall density. It needs to be continued by focusing on other structuring levels than the national level (regional and macro-regional levels) and by focusing on various disciplines.

REFERENCES

- Batagelj, V., Kejzar, N., & Korenjak-Cerne, S. (2006). Analyzing the Structure of U.S. Patents Network. In *Data Science and Classification* (p. 358). Ljubjana: Springer.
- Bornmann, L., Leydesdorff, L., Walch-Solimena, C., & Ettl, C. (2011). Mapping excellence in the geography of science: An approach based on Scopus data. *Journal of Informetrics*, 5(4), 537-546. doi:10.1016/j.joi.2011.05.005
- Eckert, D., Baron, M., & Jegou, L. (2013). Les villes et la science: apports de la spatialisation des données bibliométriques mondiales. *M@ppemonde*, 110(2). URL : <http://mappemonde.mgm.fr/num38/articles/art13201.pdf>
- Gauffriau, M., Larsen, P., Maye, I., Roulin-Perriard, A., & Ins, M. (2008). Comparisons of results of publication counting using different methods. *Scientometrics*, 77(1), 147-176. doi:10.1007/s11192-007-1934-2
- Grossetti, M., Eckert, D., Gingras, Y., Jégou, L., Larivière, V., & Milard, B. (2014). Cities and the geographical deconcentration of scientific activity: A multilevel analysis of publications (1987–2007). *Urban Studies*, 51(10), 2219-2234. doi:10.1177/0042098013506047

- Hoekman, J., Frenken, K., & Oort, F. (2009). The geography of collaborative knowledge production in Europe. *The Annals of Regional Science*, 43(3), 721-738. doi:10.1007/s00168-008-0252-9
- Inhaber, H. (1977). Changes in centralization of science. *Research Policy*, 6, 178-193.
- Leydesdorff, L., & Persson, O. (2010). Mapping the Geography of Science: Distribution Patterns and Networks of Relations among Cities and Institutes.
- Matthiessen, C. W., Schwarz, A. W., & Find, S. (2010). World Cities of Scientific Knowledge: Systems, Networks and Potential Dynamics. An Analysis Based on Bibliometric Indicators. *Urban Studies*, 9(47), 1879-1897. doi:10.1177/0042098010372683
- Royal Society (Great Britain). (2011). Knowledge, networks and nations global scientific collaboration in the 21st century. London: The Royal Society.
- Schott, T. (1993). World Science: Globalization of Institutions and Participation. *Science Technology & Human Values*, 18(2), 196-208.
- Sexton, J. (2012). State of The World's Science. *Scientific American*, 307(4), 36-40.
- Tijssen, R. J., Waltman, L., & van Eck, N. J. (2012). Research collaboration and the expanding science grid: Measuring globalization processes worldwide. arXiv preprint arXiv:1203.4194. URL : <http://arxiv.org/abs/1203.4194>
- Wagner, C. S. (2008). *The New Invisible College: Science for Development*. Brookings Institution Press.
- Zitt, M., Barré, R., Sigogneau, A., & Laville, F. (1999). Territorial concentration and evolution of science and technology activities in the European Union: a descriptive analysis. *Research Policy*, 28(5), 545-562. doi:10.1016/S0048-7333(99)00



Trends in the inter-regional and international research collaboration of the PRC's regions: 2000-2015

Marc Luwel*, Erik van Wijk*, Lambertus J. Van der Wurff* and Lili Wang**

* {luwel,wijk,wurff}@cwts.leidenuniv.nl

Centre for Science and Technology Studies (CWTS), Leiden University,

P.O. Box 905, 2300 AX Leiden (Netherlands)

**wang@merit.unu.edu

UNU-MERIT, Keizer Karelplein 19, 6211 TC Maastricht (Netherlands)

ABSTRACT

The regional structure of the PRC's scientific output is analyzed using publications processed for the Web of Science. Over the period 2000-2015 and measured by the Salton Index of the co-publications the scientific collaboration among the PRC's regions increased only slightly, in stark contrast with the USA's states and during the most recent years the EU member countries. Only for research with other nations, representing about 30% of the total publication output, inter-regional collaboration is on the rise. For the leading PRC's regions the USA is the dominant partner co-authoring about 50% of their publications. Germany and especially Japan seems to lose attractiveness to the advantage of the UK, Australia and neighboring Asian countries.

INTRODUCTION

The People's Republic of China (PRC) became the third economic power measured in nominal GDP only preceded by the USA and the European Union; the latter not being a nation it is now the second wealthiest country. In parallel its science and innovation capacity developed rapidly and measured by the number of scientific publications it is also the world's second most productive nation, again after the USA (National Science Board, 2014)

It is well known that over the last half century the business of science has become more internationally oriented and cross-border collaborations are on the rise especially for natural and life science and the basic disciplines of engineering (Waltman, 2011). Within larger countries intra-regional collaborations are also increasing (Bellini, 2013). To study these phenomena often publications are used as a proxy for a country's scientific capacity.

The PRC's international co-publications have been intensively studied (Glänzel, 2007; Haustein, 2011), often focusing on specific domains (Tang, 2011) or countries (Wagner, 2015). Far less attention has been paid to the PRC's domestic scientific collaboration and the

international collaboration patterns of its geographical regions (Liang, 2002; Scherngell, 2011; Sun, 2015; Andersson, 2014).

RESEARCH QUESTIONS, METHODOLOGY AND DATA

In the PRC there are 31 administrative mainland regions, including 23 provinces, 5 autonomous regions and 4 municipalities, further called ‘regions’.

In this paper the regions are the unit of analysis and both the evolution of inter-regional and international collaborations are studied. Trends in inter-regional collaborations are benchmarked with those of the USA and the European Union. The regions’ profiles based on the countries they are collaborating with are compared and changes over time analyzed.

Using the Web of Science (WoS), an international bibliographic database produced by Thompson Reuters that covers very well the above mentioned disciplines, information on all publications with the PRC in the byline was extracted. The search was restricted to the period 2000-2015 and to the publication types Articles, Letters and Reviews.

Generally the PRC addresses processed for the WoS contain information on the postal code of the institute’s city; the first two or three characters of these ZIP codes indicate the region. Using this information about 85% of the addresses can be assigned automatically to a region. In the remaining addresses the postal codes are missing or erroneous. A combination of manual data cleaning and algorithms based on recurring patterns in the errors allowed to increase the yield to about 95%. In this analysis the remaining 5% of the PRC’s publications that could not be assigned to a region, are not taken into account.

For each year and each region the publications are divided into two subsets: those with an address from another country in the byline (further called international co-publications) and those with only one or more PRC addresses (further called domestic publications). In the first subset publications may be signed by authors from more than one region; domestic publications can also have addresses from two or more regions, further called domestic co-publications.

For the domestic publications and for each year a symmetrical co-publication matrix is calculated with on the diagonal the number of domestic publications of each region and on the off-diagonal entries the number of co-publications between two regions. A full or integer counting scheme is used at the level of the regions assigning a co-publication fully to each contribution unit. It should be emphasized that a publication with one or more addresses from only one region is classified as a domestic publication and assigned only once to that region.

For the international co-publications a similar matrix is calculated. For each region and for each year the list of countries in the byline of the publications and their number of co-publications is extracted from the WoS; again a full counting scheme is used

To analyze the inter-regional collaboration the absolute number of publications are an indicator. It is however well known that regions’ propensity to collaborate depends on their total number of publications. Similarity measures take this effect into account (van Eck, 2009).

In this paper to quantify the collaboration strengths between the regions the Salton Index (SI) is calculated:

$$S_{ij}(t) = \frac{P_{ij}(t)}{\sqrt{P_i(t) * P_j(t)}}$$

where *i* and *j* represent the regions, $P_{ij}(t)$ the co-publications between *i* and *j* and $P_i(t)$ the number of publications of region *i*; *t* being the publication year (Luukkonen et al., 1993). The SI is a symmetrical matrix with 1 on the diagonal.

To make the evolution of the collaboration strength between the regions visible for each year the average, the median, the maximum and the minimum value of the SI is used as an indicator.

At the same time as the PRC's science and innovation system rapidly developed, other countries' inter-regional collaboration expanded. For the European Union stimulating collaboration between member states is even a long-standing policy objective. To benchmark PRC's inter-regional collaboration its SI is compared with this measure calculated for the first 15 countries joining the European Union (EU15) and for the states within the United States (Luwel, 2015).

One of the factors influencing co-publication activities among regions is not only the available scientific capacity but also their geographical locations. To test for effects of the geographical proximity on the collaborative strength between regions a symmetrical distance matrix was calculated using the geo-coordinates of the capital city of the regions. To test for correlation the Mantel test between the co-publication matrix and the distance matrix on the one hand and the SI and the latter on the other hand is done using the Pearson correlation and treating the diagonal elements in the matrices as missing values (Mantel, 1967). This operation was carried out for each year and for both subsets.

Next each region's international co-publication profile is constructed based on a country's number of co-publications with that region and on the ratio between this number and the region's total number of international co-publications. The regional profiles are compared as well as their evolution during the period 2000-2015.

RESULTS

Table 1 gives for the 31 regions the number of international co-publications and domestic publications. The distribution of these publications is highly skewed with the top-10 regions producing more than 75% of the total output. Over the 15 year period the total number of publications (i.e. the sum of the number of domestic and internationally co-authored publications) increased by a factor 10; this increase is roughly the same for the top 10 and the next 10 regions and even higher for the third tier.

In 2015 the ratio of the international co-publications and the domestic publications is about 30%, an increase by 7% compared to 2000; for the top-10 regions this increase is roughly the same.

However table 1 shows that the growth rate decreases slightly over time but the increase in internationally co-authored publications is outperforming the domestic publications in the last two 5 year periods by 44% and 25%.

Next the strength of the inter-regional collaborations and its evolution are analyzed by calculating for each year the SI. Figure 1a shows the evolution of the mean value of the SI for the international and the domestic co-publications for the 31 regions and for the top-10 regions. For the domestic co-publications the average value of the SI remains roughly constant over the period 2000-2015. It is not surprising that for all the 31 regions together the values of the indicator are lower than for the top-10 as the third tier regions have a low number of (co-)publications.

At the beginning of the period the average values of the SI for the international co-publications is below the values for the domestic publications, subsequently during the next few years they have a rather erratic behavior. In the beginning of the last decade most regions even among the top-10, had very few international co-publications; the subset of these publications with addresses from than one region was even smaller. The upswing in the values for the year 2004 could be explained by the incorporation of additional Chinese

journals in the WoS and will be investigated in more detail. From 2008-2009 onwards the average values of the SI for the international co-publications increase systematically to a level well above the values for the domestic publications. This trend is more pronounced for the top-10 regions.

For the EU15 and the USA no separate values of the SI for international and domestic co-publications are available in Luwel (2015). To benchmark the PRC figure 1b shows the average values of the SI of the two subsets together for the top-10 regions and the 31 regions and these data for the intra-EU15 co-publications and for the USA the co-publications with at least two different states.

There is a strong contrast between the USA, EU15 and the PRC. For the USA the SI average value is in 2000 already substantially higher and the indicator has a stronger growth rate during the period. Between 2000 and 2005 the SI average values for the EU15 are the same as for the top-10 regions but from 2005 onwards the increase for the EU15 is much more pronounced. For the 31 regions together during most of the period the SI average values are stable, increasing slightly only during the most recent years.

To analyze the effect of the distance between regions on their propensity to collaborate the correlation between the co-publication matrix and the SI on the hand and the matrix of the distances between the regions is calculated.

For the two subsets the co-publication matrix and the distance matrix correlate at 1 % significance level for all years except for 2001. The same result is obtained for the SI. However for the top-10 regions the Mantel test shows that the co-publication matrix and the distance matrix as well as the SI and the distance matrix are not correlated even at 5% significance level.

Each region has its own international co-publication profile which can evolve over time. In table 2 for each of the 4 regions with the most international co-publications in 2015 the top 5 countries are listed. The USA is on top and around 50% of these regions' international co-publications has at least one address from this country in the byline. However the growth rate seems to level off, especially for GuangDong. A second observation is the decrease of the fraction of the publications in collaboration with researchers working in German institutes and the inverse tendency for Australia and the UK. Another striking observation is the absence of Japan among the top 5 except for co-publications with Shanghai but its share is decreasing rapidly. The same pattern is observed for most of the top regions as is illustrated in figure 2. For the other 6 regions in the top-10 the fraction of the international co-publications of the leading countries (except the USA) is given (publication year 2015). Between 44% and 47% of these regions' international co-publications are in collaboration with this country. Australia and the UK are competing for the second place. Only for Liaoning is Japan an important partner co-singing more than 10% of its international co-publications.

DISCUSSION

The regional distribution of PRC's publications in journals processed for the WoS is highly skewed; the top-10 regions produce more than 75% of the total output. Although the growth rates are higher for the less productive regions, the share of the top-10 regions in the total regional publication output is only slowly reducing over the last 15 years. Similar skewness in the distribution of publications has been observed within the USA and between EU member states (Luwel, 2015).

The overall growth rate of both international co-publications and domestic publications slows down; comparing 2010 and 2015 this trend is more pronounced for international co-publications than for domestic publications. Their ratio is 31% in 2015. But again there is a large deviation between the top-10 regions with 33% and the third tier with 22%.

Using co-publications as a proxy in contrast with the USA and the EU the scientific collaboration between the regions is not increasing substantially. Only on international co-publications the PRC's regions are collaborating somewhat more intensely during the last couple of years. In the paper the evolution of the average value of the SI is presented; using the median value of this index and other similarity measures similar results are obtained.

For the 31 regions together the scientific collaboration measured by co-publications and geographical separation correlates very well. This result is strongly influenced by the regions with the lowest number of (co-)publications. The collaboration among the top-10 regions evolves independently from their geographical location; over the period 2000-2015 the co-publications matrix and SI are not correlated with the distance matrix.

The USA is the international partner par excellence but there are differences in regions' international co-publication profiles with the propensity to collaborate with Germany and Japan decreasing and a growing influence of Australia and the UK. Especially for Germany's and Japan's decline no obvious explanation can be given. Australia with its strong international higher education sector and neighboring Asian countries become for Chinese researchers more attractive as scientific partners.

To obtain a better understanding of the relationship between the geography and the scientific collaboration among the regions work is in progress to construct Mantel correlograms (Diniz-Filho, 2013); although they assume an underlying distribution gravity models could also be used (Scherngell, 2011). To benchmark in more detail the trends in PRC's co-publications the data for the EU and the USA will be broken down in international and domestic co-publications. Finally the analysis presented in this paper can also be made using the 10% most cited papers or for individual scientific disciplines.

REFERENCES

- Andersson D.E., Gunessee S., Matthiessen C.W. & Find S. (2014). The Geography of Chinese Science. *Environment and Planning A*, 46, 2950-2971.
- Bellini N. & Hilpert U. (Ed.). (2013). *Europe's Changing Geography: The impact of Inter-regional Networks*. Abingdon: Routledge.
- Diniz-Filho J.A.F., Soares T.N., Lima J.S., Dobrovolski R., Landeiro V.L., Pires de Campos Telles M., Rangel T.F. & Bini L.M. (2013). Mantel Test in Population Genetics. *Genetics and Molecular Biology*, 36, 475-485.
- Glänzel W., Debackere K. & Meyer M. (2008). 'Triad' or 'Tetrad'? On Global Changes in a Dynamic Worlds. *Scientometrics*, 74, 71-88.
- Haustein S., Tunger D., Heinrichs G. & Baelz G. (2011). Reasons for and Developments in International Scientific Collaboration: Does an Asia-Pacific Research Area Exits from a Bibliometric Point of View? *Scientometrics*, 86, 727-746.
- Liang L. & Zhu L. (2002). Major Factors Affecting China's Inter-regional Research Collaboration: Regional Scientific Productivity and Geographical Proximity. *Scientometrics*, 55, 287-316.
- Luukkonen T., Tijssen R.J.W., Persson O. & Sivertsen G. (1993). The Measurement of International Scientific Collaborations. *Scientometrics*, 28, 15-36.
- Luwel M & van Wijk E. (2015), Scientific Europe Coming Together? *20th International Conference on Science and Technology Indicators, Lugano, Switzerland, September 2-4, 2015*.
- National Science Board (2014). *Science and Engineering Indicators 2014*. Arlington VA: National Science Foundation (NSB 14-01).
- Mantel N. (1967). The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Research*, 27, 209-220.
- Scherngell T. & Hu Y. (2011). Collaborative Knowledge Production in China: Regional Evidence from a Gravity Model Approach. *Regional Studies*, 45, 755-772.
- Sun Y. & Cao C. (2015). Intra- and Inter-Regional Research Collaboration across Organizational Boundaries: Evolving Patterns in China. *Technological Forecasting and Social Change*, 96, 215-231.
- Tang L. & Shapira P. (2011). *China-US scientific collaboration in nanotechnology: patterns and dynamics*. *Scientometrics*, 88, 1-16.
- Van Eck, N.J., & Waltman, L. (2009). How to Normalize Cooccurrence Data? An Analysis of Some Well-Known Similarity Measures. *Journal of the American Society for Information Science and Technology*, 60, 1635-1651.
- Wagner C.S., Bornmann L. & Leydesdorff L. (2015). Recent Developments on China-US Cooperation in Science. *Minerva*, 53, 199-214.
- Waltman L., Tijssen R.J.W. & van Eck N.J. (2011). Globalisation of Science in Kilometers, *Journal of Informetrics*, 5, 574-582.

Table 1a. For the 31 regions, the number of domestic publications for the publication year 2000, 2005, 2010 and 2015 (full counting scheme). The sixth colon (%2015) gives each region's percentage of the total number of publications in 2015. The last three colons give the Percent (Straight-Line) Growth Rates for 2005-2000 (Gr 05-00), 2010-2005 (Gr 10-05) and 2015-2010 (Gr 15-10)

No	Region	2000	2005	2010	2015	% 2015	Gr 05-00	Gr 10-05	Gr 15-10
1	Prch-Beijing	6614	14925	24424	41069	17,95	125,7	63,6	68,2
2	Prch-Jiangsu	1687	4641	10678	21997	9,62	175,1	130,1	106,0
3	Prch-Shanghai	2989	7226	12737	20602	9,01	141,8	76,3	61,7
4	Prch-GuangDong	887	2487	6650	13421	5,87	180,4	167,4	101,8
5	Prch-Shandong	860	2873	5994	11840	5,18	234,1	108,6	97,5
6	Prch-Zhejiang	858	3601	6622	11316	4,95	319,7	83,9	70,9
7	Prch-Hubei	1080	3300	5963	11201	4,90	205,6	80,7	87,8
8	Prch-Shaanxi	820	2480	5481	10874	4,75	202,4	121,0	98,4
9	Prch-Sichuan	682	1929	4924	9442	4,13	182,8	155,3	91,8
10	Prch-Liaoning	948	2826	5081	8353	3,65	198,1	79,8	64,4
	<i>Total 1-10</i>	<i>17425</i>	<i>46288</i>	<i>88554</i>	<i>160115</i>	<i>70,00</i>	<i>165,6</i>	<i>91,3</i>	<i>80,8</i>
11	Prch-Tianjin	726	2198	3483	7285	3,18	202,8	58,5	109,2
12	Prch-Hunan	654	1873	3780	7069	3,09	186,4	101,8	87,0
13	Prch-Anhui	1101	2471	3587	6480	2,83	124,4	45,2	80,7
14	Prch-Jilin	875	2127	3550	6097	2,67	143,1	66,9	71,7
15	Prch-Heilongjiang	308	1408	3260	5765	2,52	357,1	131,5	76,8
16	Prch-Henan	275	963	2634	5633	2,46	250,2	173,5	113,9
17	Prch-Chongqing	185	761	2178	4592	2,01	311,4	186,2	110,8
18	Prch-Fujian	491	1304	2405	4579	2,00	165,6	84,4	90,4
19	Prch-Gansu	707	1410	2455	3680	1,61	99,4	74,1	49,9
20	Prch-Hebei	273	803	1601	3339	1,46	194,1	99,4	108,6
	<i>Total 11-20</i>	<i>5595</i>	<i>15318</i>	<i>28933</i>	<i>54519</i>	<i>23,83</i>	<i>173,8</i>	<i>88,9</i>	<i>88,4</i>
21	Prch-Jiangxi	87	417	1225	2638	1,15	379,3	193,8	115,3
22	Prch-Shanxi	179	667	1176	2365	1,03	272,6	76,3	101,1
23	Prch-Yunnan	277	575	1382	2271	0,99	107,6	140,3	64,3
24	Prch-Guangxi	104	346	977	2091	0,91	232,7	182,4	114,0
25	Prch-Xinjiang	74	189	661	1439	0,63	155,4	249,7	117,7
26	Prch-Guizhou	66	143	424	1043	0,46	116,7	196,5	146,0
27	Prch-Inner Mongolia	38	146	402	934	0,41	284,2	175,3	132,3
28	Prch-Hainan	22	67	232	624	0,27	204,5	246,3	169,0
29	Prch-Ningxia	12	43	110	335	0,15	258,3	155,8	204,5
30	Prch-Quinghai	16	78	157	307	0,13	387,5	101,3	95,5
31	Prch-Tibet	3	3	39	67	0,03	0,0	1200,0	71,8
	<i>Total 21-31</i>	<i>878</i>	<i>2674</i>	<i>6785</i>	<i>14114</i>	<i>6,17</i>	<i>204,6</i>	<i>153,7</i>	<i>108,0</i>
	Total	23898	64280	124272	228748	100,00	169,0	93,3	84,1

Table 1b. For the 31 regions, the number of international co-publications. The table has the same structure as table 1a

No	Region	2000	2005	2010	2015	% 2015	Gr 05-00	Gr 10-05	Gr 15-10
1	Prch-Beijing	1967	4103	8476	15673	22,30	108,6	106,6	84,9
2	Prch-Shanghai	687	1680	4072	8135	11,57	144,5	142,4	99,8
3	Prch-Jiangsu	468	1106	2654	6783	9,65	136,3	140,0	155,6
4	Prch-GuangDong	172	582	1799	4339	6,17	238,4	209,1	141,2
5	Prch-Hubei	257	704	1803	4064	5,78	173,9	156,1	125,4
6	Prch-Zhejiang	233	704	1837	3689	5,25	202,1	160,9	100,8
7	Prch-Shaanxi	170	453	1381	3103	4,42	166,5	204,9	124,7
8	Prch-Shandong	152	509	1259	2565	3,65	234,9	147,3	103,7
9	Prch-Sichuan	148	329	1150	2539	3,61	122,3	249,5	120,8
10	Prch-Liaoning	252	672	1245	2144	3,05	166,7	85,3	72,2
	Total 1-10	4506	10842	25676	53034	75,46	140,6	136,8	106,6
11	Prch-Anhui	246	499	1005	2054	2,92	102,8	101,4	104,4
12	Prch-Tianjin	140	430	872	1898	2,70	207,1	102,8	117,7
13	Prch-Hunan	74	308	804	1874	2,67	316,2	161,0	133,1
14	Prch-Heilongjiang	80	297	784	1595	2,27	271,3	164,0	103,4
15	Prch-Jilin	156	369	831	1422	2,02	136,5	125,2	71,1
16	Prch-Fujian	100	241	547	1404	2,00	141,0	127,0	156,7
17	Prch-Chongqing	27	119	444	1201	1,71	340,7	273,1	170,5
18	Prch-Henan	37	106	321	1042	1,48	186,5	202,8	224,6
19	Prch-Gansu	109	232	608	878	1,25	112,8	162,1	44,4
20	Prch-Yunnan	68	182	408	767	1,09	167,6	124,2	88,0
	Total 11-20	1037	2783	6624	14135	20,11	168,4	138,0	113,4
21	Prch-Hebei	33	116	269	556	0,79	251,5	131,9	106,7
22	Prch-Shanxi	36	117	163	556	0,79	225,0	39,3	241,1
23	Prch-Jiangxi	11	57	177	499	0,71	418,2	210,5	181,9
24	Prch-Guangxi	25	89	195	432	0,61	256,0	119,1	121,5
25	Prch-Xinjiang	16	39	139	321	0,46	143,8	256,4	130,9
26	Prch-Guizhou	24	44	111	280	0,40	83,3	152,3	152,3
27	Prch-Inner Mongolia	17	24	102	154	0,22	41,2	325,0	51,0
28	Prch-Hainan	2	8	44	144	0,20	300,0	450,0	227,3
29	Prch-Quinghai	2	20	40	88	0,13	900,0	100,0	120,0
30	Prch-Ningxia	1	10	30	60	0,09	900,0	200,0	100,0
31	Prch-Tibet	4	3	15	23	0,03	-25,0	400,0	53,3
	Total 21-31	171	527	1285	3113	4,43	208,2	143,8	142,3
	Total	5714	14152	33585	70282	100,00	147,7	137,3	109,3

Table 2. For the first 4 regions in the top-10 the 5 countries with the highest number of co-publications are listed as well as the ratio of this number and the region's total number of international co-publications for each uneven year between 2000 and 2015

Region	Country	2001	2003	2005	2007	2009	2011	2013	2015
Beijing	USA	0,40	0,43	0,47	0,46	0,46	0,48	0,51	0,52
Beijing	UK	0,09	0,09	0,12	0,10	0,10	0,10	0,12	0,12
Beijing	Germany	0,13	0,13	0,12	0,10	0,09	0,10	0,11	0,10
Beijing	Australia	0,04	0,05	0,05	0,07	0,06	0,07	0,08	0,09
Beijing	Canada	0,05	0,05	0,08	0,07	0,08	0,07	0,07	0,08
Jiangsu	USA	0,35	0,33	0,38	0,38	0,45	0,45	0,46	0,48
Jiangsu	Australia	0,05	0,06	0,06	0,07	0,06	0,09	0,09	0,09
Jiangsu	UK	0,06	0,07	0,06	0,08	0,09	0,09	0,08	0,09
Jiangsu	Canada	0,06	0,05	0,07	0,07	0,08	0,07	0,07	0,07
Jiangsu	Germany	0,06	0,10	0,08	0,08	0,08	0,07	0,07	0,07
Shanghai	USA	0,34	0,38	0,43	0,43	0,45	0,48	0,53	0,53
Shanghai	UK	0,08	0,08	0,09	0,08	0,10	0,09	0,10	0,11
Shanghai	Australia	0,06	0,06	0,03	0,05	0,06	0,06	0,08	0,09
Shanghai	Germany	0,12	0,10	0,08	0,08	0,08	0,08	0,08	0,09
Shanghai	Japan	0,24	0,20	0,18	0,15	0,12	0,10	0,09	0,09
GuangDong	USA	0,31	0,33	0,29	0,38	0,47	0,48	0,51	0,51
GuangDong	Australia	0,05	0,09	0,13	0,12	0,09	0,09	0,10	0,13
GuangDong	UK	0,07	0,07	0,08	0,08	0,08	0,09	0,08	0,11
GuangDong	Canada	0,06	0,08	0,09	0,08	0,06	0,07	0,07	0,08
GuangDong	Germany	0,19	0,11	0,08	0,07	0,05	0,07	0,07	0,08

Figure 1a. Evolution of the SI average value of the international co-publications (a) and the domestic co-publications (b) of the 31 regions and of the top-10 regions (c and d)

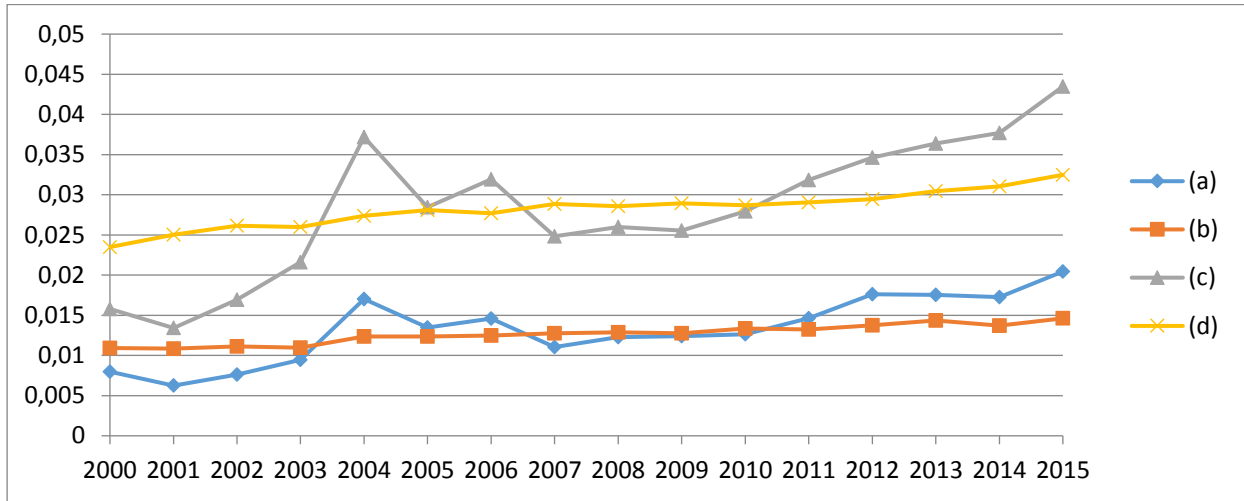


Figure 1b. Evolution of the SI average value of the co-publications of the 31 regions (a), the top-10 regions (b), the EU15 (c) and the USA (d)

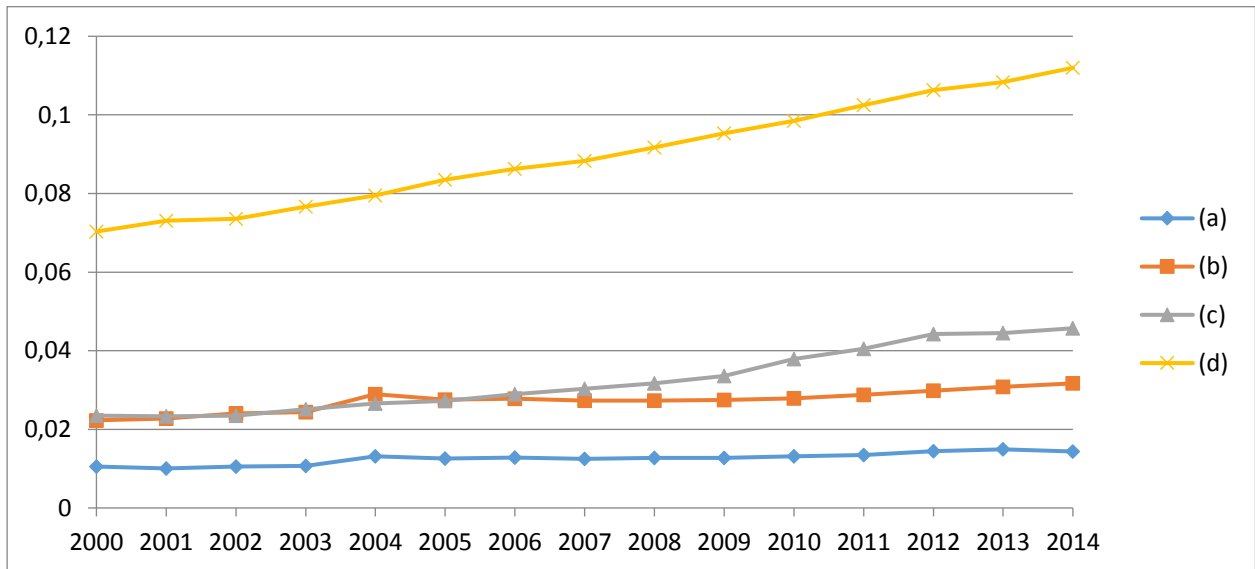
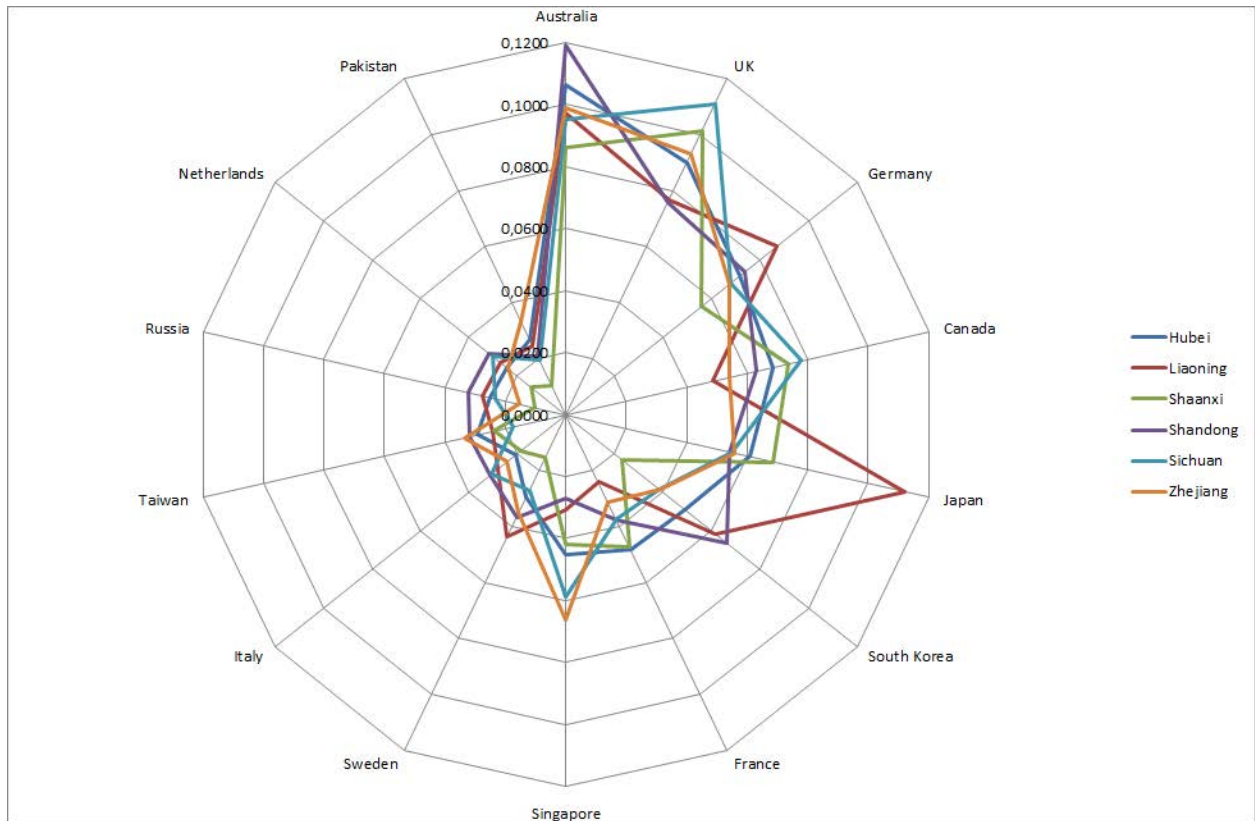


Figure 2. For Hubei, Liaoning, Shaanxi, Shandong, Sichuan and Zhejiang the countries with the most co-publications in 2015 (after the USA, not shown in the graph) are given as well as the ratio of this number and the region’s total number of international co-publications.





Iran's scientific dominance and the emergence of South-East Asian countries in the Arab Gulf Region

Henk F. Moed*

*Independent researcher and senior scientific advisor, Amsterdam, The Netherlands.
Email: hf.moed@gmail.com

ABSTRACT

A longitudinal bibliometric analysis of publications indexed in Thomson Reuters' Incites and Elsevier's Scopus, and published from the Arab Gulf States and neighbouring countries, shows clear effects of major political events during the past 35 years. Predictions made in 2006 by the US diplomat Richard N. Haass on political changes in the Middle East have come true in the Gulf States' national scientific research systems, to the extent that Iran has become in 2015 by far the leading country in the Arab Gulf, and South-East Asian countries including China, Malaysia and South Korea have become major scientific collaborators, displacing the USA and other large Western countries. But collaborations patterns among Gulf States show no apparent relationship with differences in Islam denominations.

INTRODUCTION

Political developments in the Arab Gulf Region are still in the centre of global public interest. A commentary published in Nature in 1991 shortly after the start of Operation Desert Storm analysed the scientific integration of 12 Arabic Gulf States and the western world during the 1980s, and compared these patterns with changes in international political relations (De Bruin, Braam and Moed, 1991). Its base assumptions stated that international scientific collaboration patterns reflect geographical, political, social and historical relations (Frame and Carpenter, 1979) and that it is important for all stakeholders to have a thorough understanding of the relationships in an area of political tension.

This short communication provides an update of the 1991 study by De Bruin, Braam and Moed. It presents a longitudinal bibliometric analysis of publications published from the Gulf States and neighbouring countries and indexed in Thomson Reuters' Web of Science and Elsevier's Scopus, covering a time period as long as 35 years (1980-2015). The US diplomat Richard N. Haass predicted in 2006 trends in what he termed the upcoming "Middle East fifth era" (Haass, 2006). This paper empirically examines four of Haass' key predictions, namely that "the United States will continue to enjoy more influence in the region than any other outside power, but its influence will be reduced from what it once was"; that "United States will increasingly be challenged by the foreign policies of other outsiders"; Iran will be one of the two most powerful states in the region"; and "tensions between Sunnis and Shiites will grow throughout the Middle East".

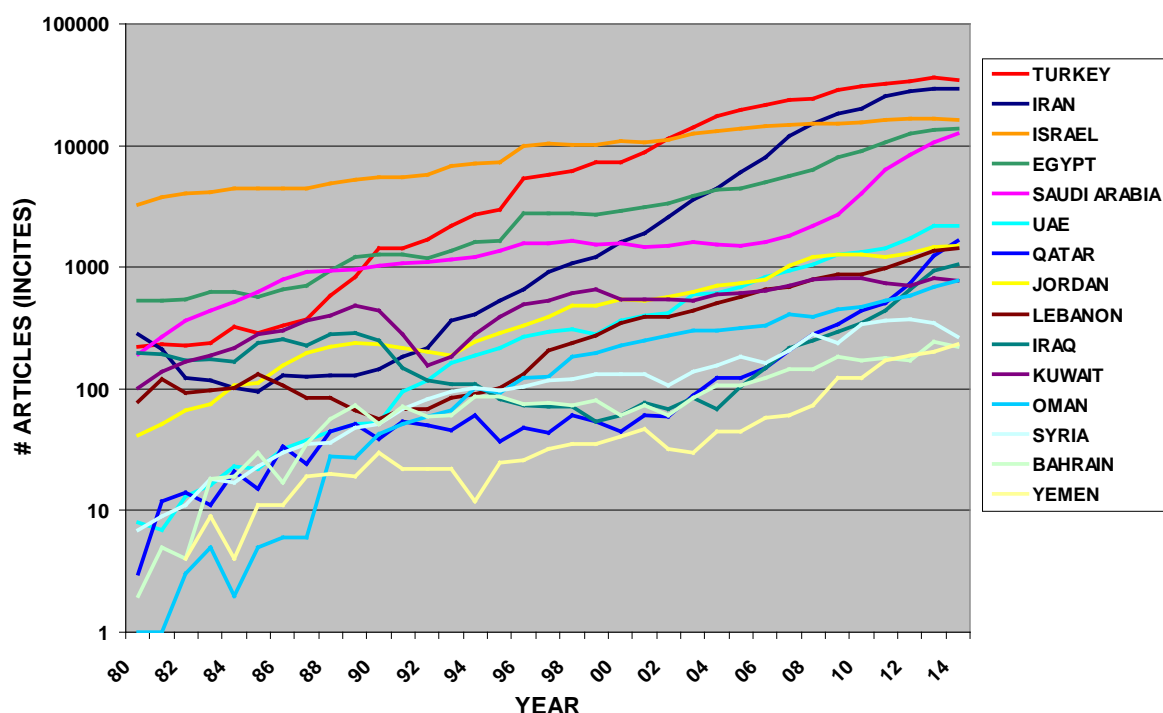
RESULTS

Trends in the number of publications and internationally co-authored publications

Figure 1 shows that Qatar and United Arab Emirates (UAE) had the largest increase in publication counts during 1980-2014: three orders of magnitude, and Kuwait and Iraq the smallest (zero and one order of magnitude, respectively). The latter outcome suggests that Kuwait has never overcome the devastations of the 1990-1991 Persian Gulf War. Iraq recuperated to some extent after the 2003 invasion by a United States-led coalition (Operation Iraqi Freedom) and the deposition of the Ba'athist government of Saddam Hussein, as it started revealing moderate positive annual growth rates as from 2005.

Iran and Saudi Arabia had in 2014 the largest absolute number of publications of all Gulf States, namely 29,000 and 13,000, respectively. Iran's research output declined during the first half of the 1980s under the influence of the Iraq-Iran War which started in September 1980, but as from the beginning of the 1990s, when the Persian Gulf War started with Iraq's invasion of Kuwait, it revealed an exponential growth, doubling up until 2011 approximately every 3 years. Saudi Arabia's annual publication counts showed almost flat growth rates during 1990-2007, but rapidly increased to around 50 per cent in 2010-2011, but then declined to 20 per cent in 2013-2014.

Figure 1: Annual number of publications during 1980-2014 for 12 Gulf States and three large neighbouring countries



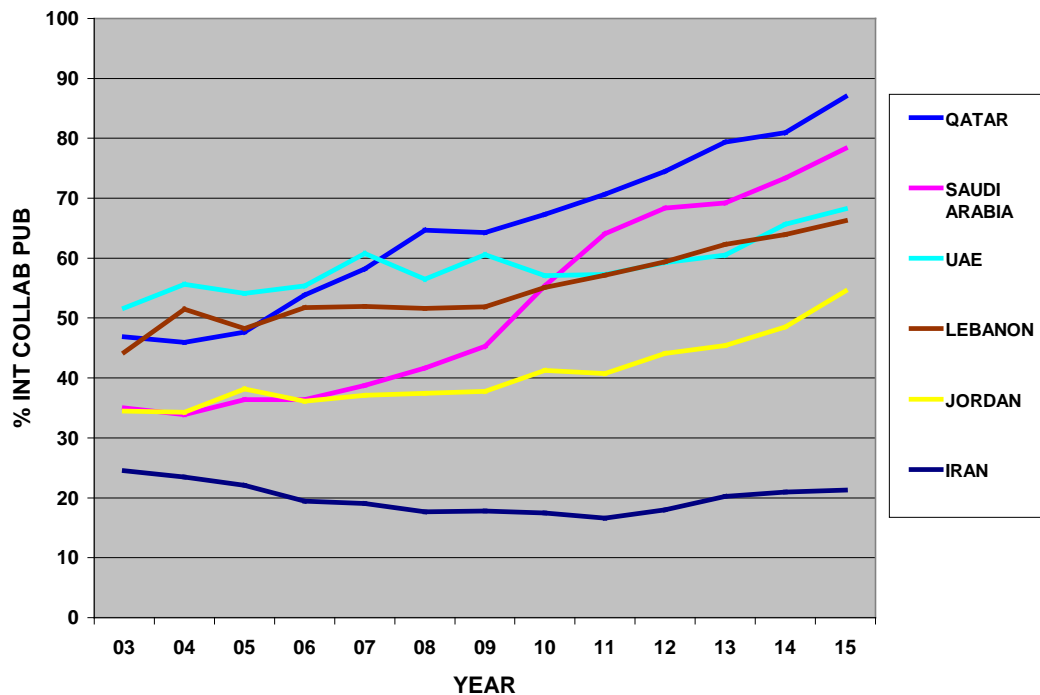
Legend to Figure 1. Data were extracted from Thomson Reuters' Incites. Since data for 2015 are not yet complete they not displayed. The number of 2015-articles indexed up until 5 December 2015 amounts to 20,400 for Iran against 17,400 for Turkey.

According to data were extracted from Scopus, using its subject classification into 27 disciplines, for USA, UK and other larger Western countries medicine tends to be the most important discipline, with typically 20 percent of publications, followed by engineering and biochemistry, genetics & molecular biology, each with some 10 per cent. But in China and Malaysia these percentages are reversed, while India, Pakistan and South Korea have an

intermediary position. The latter is also true for most Gulf States. The share of their papers in medicine and in biochemistry, genetics & molecular biology is in most cases lower than that of larger western countries, and that in engineering, physics & astronomy and chemistry higher. Exceptions are Bahrain and Lebanon, with 28 and 24 percent of articles in medicine, respectively.

Large differences exist between Iran and Saudi Arabia with respect to the amount of foreign input needed to produce these papers. According to [Figure 2](#), the percentage of internationally co-authored publications (ICAP) is in 2015 almost 80 per cent for Saudi Arabia, but only around 20 per cent for Iran. Applying a model of scientific development presented in [Table 1](#), the results suggest that the two countries are in different phases of scientific development. While Saudi Arabia and most other Gulf States are still in the building up phase, Iran is currently moving from a consolidation and expansion into the internationalization phase. This trend can be expected to continue now that the international boycotts are cancelled. The other Gulf States still depend in various degrees upon collaboration with external institutions, increase their ICAP rate and are, in terms of the scientific development model presented in [Table 1](#), in a phase of building up a scientific infrastructure. This is especially true for the two countries showing the largest increase of their publication output, namely Qatar and UAE, with ICAP percentages of 87 and 68, respectively.

Figure 2. Percentage of internationally co-authored publications (relative to total publication output) of major Gulf States (Data from Elsevier's Scopus)



Gringas (2014) found that a disproportionately large number of researchers appearing in the Thomson Reuters' list of highly cited researchers indicate Saudi universities as secondary address, thus boosting these institutions up in global university rankings. The articles in which this occurs are counted as internationally co-authored publications in [Figure 2](#) and can be expected to boost up the Saudi percentage of ICAP as well. Even if the influence of this phenomenon is substantial, it underlines the dependence of this country upon the input of foreign researchers and does therefore not violate the conclusions on its state of scientific

development. More bibliometric data on Middle East countries derived from Incites can be found in Gul et al. (2015).

Table 1. A bibliometric model for capturing the state of a country's scientific development

Phase	Description	Trend in # published articles	Trend in % internationally co-authored publications
Pre-development	Low research activity without clear policy of structural funding of research	~	~
Building up	Collaborations with developed countries are established; national researchers enter international scientific networks	+	++
Consolidation and expansion	The country develops its own infrastructure; the amount of funds available for research increases	++	-
Internationalization	Research institutions in the country start as fully-fledged partners, increasingly take the lead in international collaboration	+	+

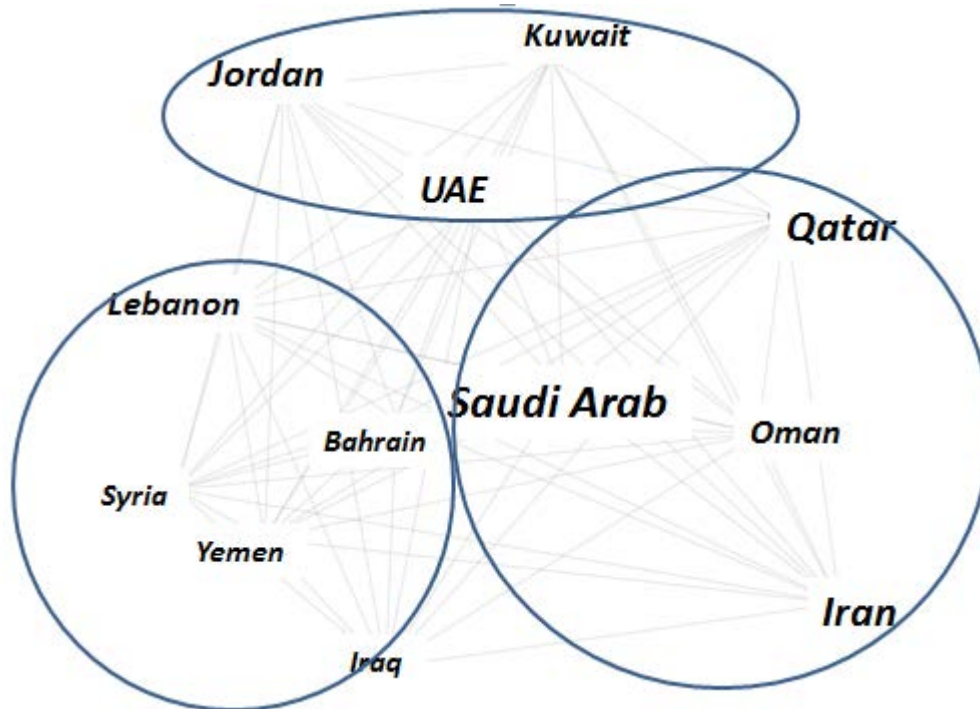
Legend to Table 1. ~ denotes: no clear trend; +: increase; -: decline; ++: strong increase. Source: UNESCO (2014). For more information on this model, see Moed and Halevi (2014).

Haass' third prediction states that "Iran will be one of the two most powerful states in the region" (Haass, 2006). During the past 3 decades, the only country that has been able to create and expand a research infrastructure of its own is Iran, despite the economic boycotts to which it has been subjected during most of the time. In terms of scientific development, Iran is clearly the leading country in the Gulf region. As from 2007, Iran's annual count exceeded both that of Egypt and Israel, and in 2015, – based on an analysis of about 50 per cent of the total number of 2015-articles eventually published –, also Turkey.

Figure 3 presents a VOS map of scientific collaborations among the 12 Gulf States. Similar to Multi-Dimensional Scaling (MDS), VOS aims to locate items in a low-dimensional space in such a way that the distance between any two items reflects the similarity of the items as accurately as possible, but differs from MDS in the way in which it attempts to achieve this aim (Van Eck, Waltman, Dekker & Van den Berg, 2010). The clustering model is a variant of modularity based clustering, a technique aiming to maximize a modularity measure of a network, defined as the fraction of the links that fall within a given group minus the expected such fraction if links were distributed at random. The VOS technique is a weighted and parameterized variant able to detect small clusters or communities (Waltman, Van Eck & Noyons, 2010). Located at the left hand side is a community with five countries with Shia dominance either within a country's Muslim population or in its government (Syria and Yemen). In the remaining two clusters all countries but one have a Sunni dominance. The striking exception is Iran in the right hand cluster, in which 90 per cent of population is Shia.

Collaboration among Gulf States and between Gulf States with countries outside the region

Figure 3. VOS Viewer map of the international co-authorship relations among all 12 Gulf States for the year 2015.

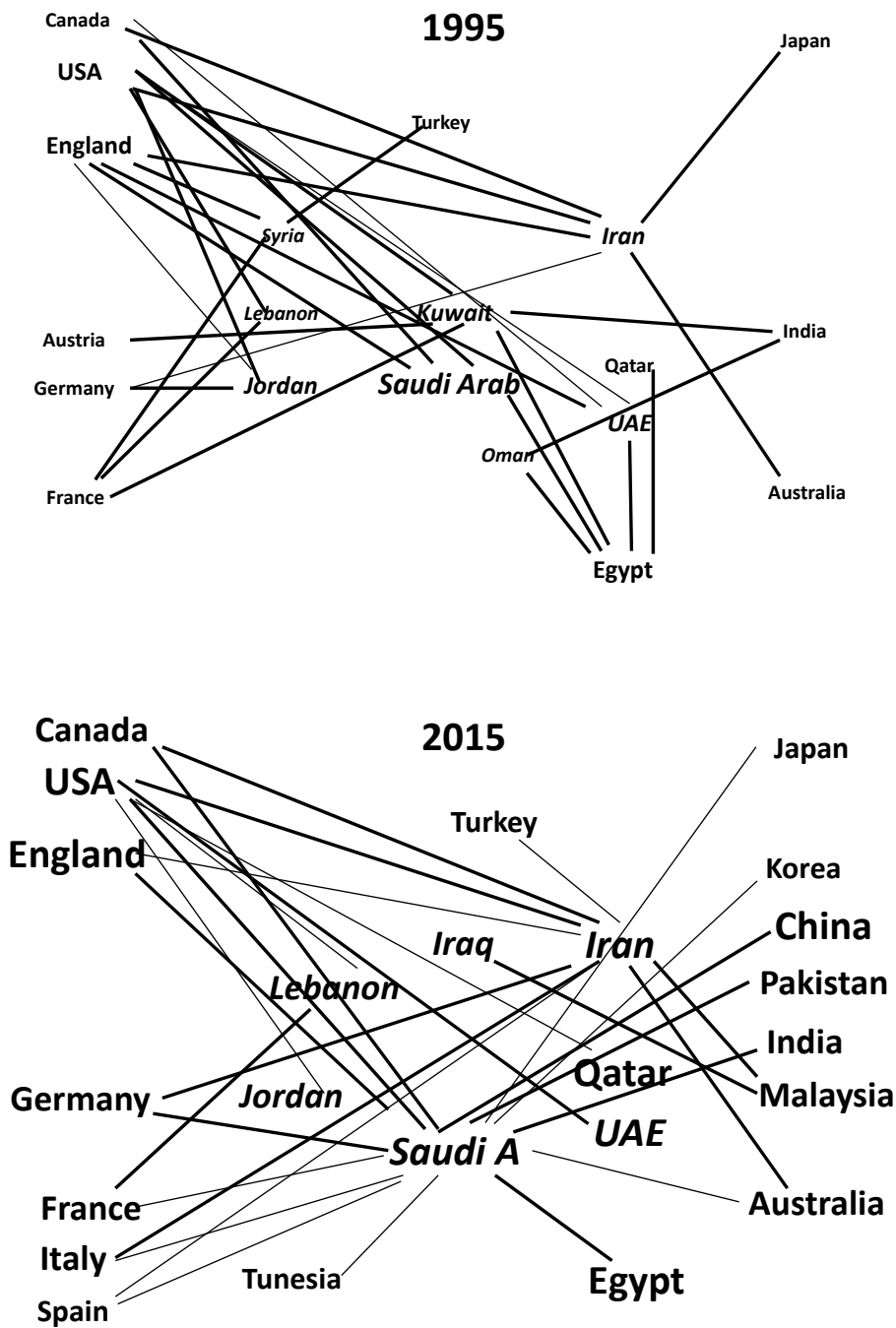


Legend to Figure 3. Data were extracted from Thomson Reuters' Incites. The three circles are inserted by the author of this paper and indicate clusters.

The community structure in the international co-authorship network among the 12 Gulf States displayed in [Figure 3](#) can be partially interpreted as traces of the main dominations within the Islam, with United Arab Emirates, Lebanon and Saudi Arabia as bridges between the two. But Iran's striking appearance in 2015 in a cluster of countries in which the majority of the population is Sunni ("Islam", n.d.) does not seem to align with Haass' assertion that "tensions between Sunnis and Shiites will grow throughout the Middle East". Intensifying collaborations with Sunny dominated countries does align with a strategy by Iran aiming to become a member of the Arab League, a regional organization of currently 22 Arab countries to which Iran has applied for membership ("Arab League", n.d.).

Analyzing collaborations between Gulf States and countries outside the region ([Figure 4](#)), the most striking feature is the emergence of East and South Asian countries during the past decade, namely China and South Korea in the East, Malaysia in the Southeast and Pakistan and India in the South. Malaysia has in 2015 strong links both with Iran and Iraq, while the other four South-East Asian countries have links with Saudi Arabia only. Iran shows a strong orientation towards Northern America and Western Europe; Malaysia is the only South-East Asian country with which Iran is linked in the map. Saudi Arabia shows a more balanced position towards Western and Asian countries as it has links with five South-East Asian countries, but also with 7 Western countries. The two emerging countries Qatar and UAE have the strongest ties with the USA.

Figure 4. Collaboration ties (ICAP) between Gulf States and countries outside the region for the years 1995 and 2015



Legend to Figure 4. Data are extracted from Incites. Figures are constructed in the same manner as those presented in De Bruin, Braam and Moed (1991) for the 1980s. They show the 30 strongest links in a particular year. Countries are positioned in a topological map. Font size indicates the number of co-authorship links, by grouping countries into quartiles on the basis of the number of co-publications in any year. The thickness of the lines the strength of the collaboration (Salton's index), applying a similar quartile approach.

Co-authorship links between Gulf States and countries outside the region are typically one order of magnitude stronger than those among Gulf States. Russia does not appear in Figure 4 or Table 2. The overall level of international scientific collaboration in this country is still low, due to historical factors, although it is increasing (Kotsemir et al., 2015). In its collaborations with the Gulf States, Russia switched from Syria in 1995 to Saudi Arabia and Iran in 2015.

Table 2 shows that in 2015 the USA is still the most important external scientific partner in the Gulf. But compared to the situation in 2005, the strength of the average co-authorship relation between USA and the 12 Gulf States declined with 25 per cent. Four South-East Asian countries, Malaysia, China, Pakistan and South Korea show large positive growth rates. While before 2008 China's ties with Iran and Saudi Arabia were of similar strength, in 2011 China started showing a preference for the latter country; in 2015 the strength of their ties is more than twice that between China and Iran. The tie between China and Saudi Arabia is among the three strongest in the region in 2015. Only the ties of the latter country with Egypt and with USA are stronger. In 2009, China became the largest importer of oil from the Gulf, but also surpassed the United States as the largest single exporter to the region as well. China's main oil provider is actually Saudi Arabia (Wakefield & Levenstein, 2011). Table 3 shows with which external countries Gulf States preferred to collaborate in the various disciplines in 2014.

Table2: Co-authorship strength between non-Gulf and Gulf countries

Rank	Country	Mean co-authorship strength in 2015	% Change compared to 2005
1	USA	0.077	-25 %
2	England	0.052	-34 %
3	Malaysia	0.050	+138 %
4	France	0.048	-26 %
5	Egypt	0.048	-19 %
6	Germany	0.044	-26 %
7	Canada	0.043	-31 %
8	China	0.039	+48 %
9	India	0.039	-2 %
10	Italy	0.037	+2 %
11	Australia	0.036	-2 %
12	Turkey	0.033	-8 %
13	Spain	0.032	+8 %
14	Pakistan	0.031	+32 %
15	South Korea	0.030	+106 %

Legend to Table 2. Co-authorship strength between two countries is defined as the number of co-authorship links between them weighted on their total number of co-authorship links (Salton's Index). Underlying data were extracted from Thomson Reuters' Incites.

The outcomes presented in Figure 4 and Table 2 fully align with Haass' first and second trend, namely that "the United States will continue to enjoy more influence in the region than any other outside power, but its influence will be reduced from what it once was", and that "United States will increasingly be challenged by the foreign policies of other outsiders". Other large Western countries show a decline as well. Instead, Malaysia, China, Pakistan and South Korea have substantially increased the collaboration in the Gulf during the past 10 years. Malaysia is now even the third country partner in the Gulf.

Table 3. Preferred foreign collaborators of Gulf States by discipline in 2014.

Gulf State	Collab country	discipline	Gulf State	Collab country	discipline
IRAN	Canada	Computer Science	SAUDI ARABIA	Canada	Computer Sci Medicine
	Germany	Physics & Astron		China	Mathematics
	Italy	Physics & Astron		France	Physics & Astron
	Malaysia	Environmental Sci		Germany	Agr & Biol Sci Physics & Astron
	Spain	Agr & Biol Sci		India	Pharmacol, Toxicol
		Physics & Astron		Italy	Physics & Astron
	Turkey	Mathematics		Japan	Physics & Astron
		Physics & Astron		S Korea	Chemical Eng Materials Science Physics & Astron
UK	Physics & Astron	Spain			Mathematics Physics & Astron
JORDAN	USA			Medicine	Tunesia
LEBANON	France	Chemistry			
		Physics & Astron			
	USA	Medicine			
QATAR	USA	Physics & Astron			

Legend to Table 3. Data were extracted from Scopus, using its subject classification into 27 disciplines, and relate to the year 2014. All collaboration pairs displayed in Figure 4 (year 2015) were analyzed. Per discipline a ratio was calculated of the percentage of internationally co-authored papers between a Gulf State and a foreign country and this gulf state's overall percentage of co-authorships with any country. Table 3 includes only cases for which this ratio exceeds 1.5.

CONCLUDING REMARK

International scientific collaboration patterns reflect geographical, political, social and historical relations and may also actively contribute to shaping these, and thus have an effect upon political relations as well. The recently established bilateral trade agreement of \$600 billion between China and Iran in the next decade (Sharafedin, 2016) illustrates how Iran's dominance and the increasing role of main South-East Asian countries in the scientific development of the Gulf States have a clear correlate in the political domain.

ACKNOWLEDGEMENT

The author wishes to thank Thomson Reuters for providing access to Incites, and the Sapienza University of Rome for offering its library facilities, including access to Elsevier's Scopus.com.

REFERENCES

- Arab League (n.d.). From https://en.wikipedia.org/wiki/Arab_League-Iran_relation.
- De Bruin, R.E., Braam, R.R., Moed, H.F. (1991). Bibliometric lines in the sand. *Nature* **349**, 559-562.
- Frame, J.D. & Carpenter, M. P. (1979). International research collaboration. *Social Studies of Science*, 9, 481-497.
- Gringas, Y. (2014). How to boost your university up the rankings. *University World News*. 18 July 2014 Issue No:329.
<http://www.universityworldnews.com/article.php?story=20140715142345754>
- Gul, S., Nisa, N., Shah, T.A., Gupta, S., Jan, A., Ahmad, S. (2015). Middle East: Research productivity and performance across nations. *Scientometrics*. 105:1157–1166.
- Haass, R.N. (2006). The New Middle East. *Foreign Affairs*.
<https://www.foreignaffairs.com/articles/middle-east/2006-11-01/new-middle-east>.
- Islam (n.d.). Information was obtained from a series of articles available at https://en.wikipedia.org/wiki/Islam_in_X, X being the name of the country.
- Kotsemir M., Kuznetsova T., Nasybulina E., Pikalova A. (2015) Identifying Directions for the Russia's Science and Technology Cooperation. *Foresight and STI Governance*, 9, 54–72.
- UNESCO (2014). *Higher Education in Asia: Expanding Out, Expanding Up*. ISBN 978-92-9189-147-4 licensed under CC-BY-SA 3.0 IGO. Montreal: UIS.
<http://www.uis.unesco.org>.
- Moed, H.F. and Halevi, G. (2014). Tracking scientific development and collaborations – The case of 25 Asian countries. *Research Trends*, Issue 38, September 2014.
- Sharafedin, B. (2016). Iran's leader says never trusted the West, seeks closer ties with China. *Markets*, Sat 23 January, 2016.
<http://www.reuters.com/article/us-iran-china-idUSKCN0V109V>.
- Van Eck, N.J., Waltman, L., Dekker, R., & Van den Berg, J. (2010). A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS. *Journal of the American Society for Information Science and Technology*, 61(12), 2405-2416.
- Wakefield, B., Levenstein, L. (eds.) (2011). *China and the Persian Gulf: Implications for the United States*. Woodrow Wilson International Center for Scholars, Washington, D.C..
- Waltman, L., Van Eck, N.J., & Noyons, E.C.M. (2010). A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics*, 4(4), 629-635.



How international is internationally collaborated research? A bibliometric study of Russian surname holder collaboration networks¹

Maria Karaulova*^{1,2}, Abdullah Gök**¹ and Philip Shapira**^{1,3}

*mkaraulova@gmail.com

¹ Manchester Institute of Innovation Research (MioIR), University of Manchester, Oxford Road, Manchester, UK, M13 9PL

² Institute of Management, Scuola Superiore Sant'Anna, Piazza Martiri della Libertà, Pisa, Italy, 56127

³ School of Public Policy, Georgia Institute of Technology (GaTech), 685 Cherry St, Atlanta, USA, 30332-0345

** abdullah.gok@mbs.ac.uk; pshapira@mbs.ac.uk

INTRODUCTION

International research performance indicators attain increased attention in science policy. They are seen to reflect relative competitiveness of a country in producing leading research (in terms of cited papers) and its commercialisation (in terms of assigned patents). However, more studies point to ongoing global bias in production, composition and assessment of research performance metrics (Rafols et al., 2012; van Leeuwen et al., 2001). As research performance indicators are used increasingly in national science policy and in influential international rankings, it is important to understand their inherent bias.

For instance, explosive growth of international collaboration in science is widely reported (Glänzel, 2001), and is generally perceived as having beneficial 'knowledge exchange' effect for involved parties. It is recognised as a capacity-building factor of domestic research indicating the increase in research quality (Bornmann et al., 2015). However, existing research has reported reproduction of uneven global relations between countries in terms of science and technology. For example, patterns of international cooperation in nanotechnology are still centred on the developed countries, which are key nodes in international networks (Shapira and Wang, 2010).

International collaboration as an indicator is routinely operationalised as co-authorship of articles between organisations by taking organisational address as a proxy of the collaborating country (Katz and Martin, 1997). For instance, many studies consider all co-publications between authors whose addresses are in different countries as international collaboration, without any attention to the backgrounds and/or previous affiliations of authors. We challenge this assumption by examining authorship data of internationally co-authored publications.

In the preliminary analysis, by using geographical approximation of author heritage rooted in the morphology of the surname, we find that in a significant minority of internationally collaborated papers, co-authors are likely to have the same origin. In other words, we observe

¹ This work was supported by the Economic and Social Research Council [grant number ES/J012785/1] as part of the project Emerging Technologies, Trajectories and Implications of Next Generation Innovation Systems Development in China and Russia.

an overestimation in the international collaboration indicator. We then investigate the dynamics of this bias by exploring conditions within which the bias changes.

By unpacking the inherent international collaboration bias, we, therefore, question the assumed relationship between co-authors in established international collaboration metrics. Ultimately, international collaboration indicators may point to reproduction and reinforcement of relationships between global centres and peripheries, and to knowledge channelling, rather to knowledge exchange.

METHODOLOGY

Surname data has been used in bibliometric analyses to determine contribution of recognisable ethnic groups to the development of particular discipline (Kissin, 2011), to determine effects of inter-ethnic collaboration on quality of publications (Freeman and Huang, 2014), or to highlight the contribution of ethnic and gender minorities (Lewison, 2001). This paper employs surname approach to identify ethnic (and national) groups that reside outside of their national borders.

The empirical focus is out-of-country collaborations of Russian addressed authors publishing in nanotechnology in the Web of Science outlets. Russia is a country that experienced growth of international collaboration rates after the breakout from Soviet isolation but also went through large-scale brain drain of its best scientists (Graham and Dezhina, 2008). Taking Russia as a case study of this research has another benefit: in a country that was internationally isolated for the large part of the 20th century, geographical approximation of ‘Russian’ surnames is consistent with the actual population, i.e. most Russians still live in Russia (Revazov et al., 1986).

This research uses nanotechnology as a basis of inquiry. Lexicological nanotechnology search query (Arora et al., 2013) is used to identify publications with author addresses in Russia in 1990-2012. Authorship structure of internationally collaborated publications is used to analyse the contribution of science diaspora to internationalisation of the Russian science. A two-step query based on morphology of typical Russian surnames (Unbegaun, 1972) is developed to identify Russian authors abroad. The resulting data is compared with the pattern of international collaborations of domestic researchers in Russia (Karaulova et al., 2016).

The dataset includes 33,538 publications that have at least one Russian addressed author. Within the dataset 93.2% of publications have a co-author and 46.2% of those collaborated papers have a co-author with an address outside of Russia. Using the surname-based lexicological method, each author in the dataset is marked either as Russian heritage or non-Russian heritage.

INITIAL FINDINGS

Initial findings from the pilot set of publications of 2010-2012 indicate a complex collaboration pattern between Russia and Russian heritage (diaspora) authors.

Countries that have extensive international collaboration networks and are the ‘core’ of science globalisation, such as the USA (Wagner and Leydesdorff, 2005), have lower rates of overseas diaspora involvement in the structure of collaboration networks with Russia (Table 1).

Table 1 Russian Diaspora Collaboration Rates in Nanotechnology

Collaborator Country	Publications of Russia Collaborated Internationally	Among them, Publications co-authored with Russian heritage authors	Share of foreign addressed collaborations with Russian heritage authors
Germany	814	190	23.34%
USA	663	226	34.09%
France	425	93	21.88%
UK	254	104	40.94%
Poland	215	27	12.56%
Japan	176	42	23.86%
Spain	157	53	33.76%
Italy	147	18	12.24%
Sweden	138	28	20.29%
Finland	135	45	33.33%

While the share of research papers collaborated with the participation of Russian heritage authors reaches 40% for major international partners of Russia, the results are more telling for minor partners. Countries that have relatively strong science base, but do not have traditionally close links with Russia, such as Portugal, Belgium or Australia, demonstrate very high level of overseas diaspora involvement in the share of publications co-authored with Russian scientists.

If a significant share of international collaborations of a national research system occurs with researchers previously affiliated with this system, internationally co-authored publications are therefore only ‘inter-national’ on a formal inter-organisational level, but in fact occur between co-authors that share academic upbringing and culture. This finding contributes to the evidence that stresses more complex nature of scientific collaboration (Bozeman and Corley, 2004) and may have fundamental implications on the use of international collaboration indicator and on science policy decisions.

Supposedly, the inherent bias in the established international collaboration indicator therefore overestimates the impact of international collaboration on periphery countries in comparison with its impact on advanced core countries. This paper makes a call for revision and further detalisation of the indicator that is sensitive to unequal science development dynamics.

CONCLUSION

When bibliometric tools are used to measure international collaboration and cooperation, invariably, assumptions are made about the social reality of these tools. Globally, the findings of this study are valid for national science policy of countries that rely on international collaboration networks to foster the development of domestic science and technology through knowledge transfer and spillovers.

Furthermore, the distribution pattern of inter-national collaboration of scientists is heterogeneous and may change over the years. The distribution of papers collaborated internationally between Russian authors may be skewed towards certain outlets and have different

patterns in different countries beyond initial findings presented above. Finally, it is important to not underestimate the role of third countries as mediators or competitors for global collaboration links.

REFERENCES

- Arora, S.K., Youtie, J., Carley, S., Porter, A.L., Shapira, P., 2013. Measuring the development of a common scientific lexicon in nanotechnology. *Journal of Nanoparticle Research* 16, UNSP 2194. doi:10.1007/s11051-013-2194-0
- Bornmann, L., Wagner, C., Leydesdorff, L., 2015. BRICS countries and scientific excellence: A bibliometric analysis of most frequently cited papers. *Journal of the Association for Information Science and Technology* 66, 1507–1513.
- Bozeman, B., Corley, E., 2004. Scientists' collaboration strategies: implications for scientific and technical human capital. *Res. Policy* 33, 599–616. doi:10.1016/j.respol.2004.01.008
- Freeman, R.B., Huang, W., 2014. Collaborating With People Like Me: Ethnic co-authorship within the US (Working Paper No. 19905). National Bureau of Economic Research.
- Glänzel, W., 2001. National characteristics in international scientific co-authorship relations. *Scientometrics* 51, 69–115. doi:10.1023/A:1010512628145
- Graham, L.R., Dezhina, I., 2008. *Science in the New Russia: Crisis, Aid, Reform*. Indiana University Press.
- Karaulova, M., Gok, A., Shackleton, O., Shapira, P., 2016. Nanotechnology Research in Post-Soviet Russia: Science System Path-Dependencies And Their Influences. *Scientometrics* Forthcoming.
- Katz, J.S., Martin, B.R., 1997. What is research collaboration? *Res. Policy* 26, 1–18. doi:10.1016/S0048-7333(96)00917-1
- Kissin, I., 2011. A surname-based bibliometric indicator: publications in biomedical journal. *Scientometrics* 89, 273–280. doi:10.1007/s11192-011-0437-3
- Lewis, G., 2001. The quantity and quality of female researchers: A bibliometric study of Iceland. *Scientometrics* 52, 29–43. doi:10.1023/A:1012794810883
- Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., Stirling, A., 2012. How journal rankings can suppress interdisciplinary research: A comparison between *Innovation Studies* and *Business & Management*. *Res. Policy* 41, 1262–1282. doi:10.1016/j.respol.2012.03.015
- Revazov, A., Paradeeva, G., Rusakova, G., 1986. [Suitability of Russian surnames as a “quasi-genetic” marker]. *Genetika* 22, 699–704.
- Shapira, P., Wang, J., 2010. Follow the money. *Nature* 468, 627–628. doi:10.1038/468627a
- Unbegaun, B.O., 1972. *Russian surnames* / B.O. Unbegaun. Clarendon Press, Oxford.
- van Leeuwen, T.N., Moed, H.F., Tijssen, R.J.W., Visser, M.S., van Raan, A.F.J., 2001. Language biases in the coverage of the Science Citation Index and its consequences

for international comparisons of national research performance. *Scientometrics* 51, 335–346. doi:10.1023/A:1010549719484

Wagner, C.S., Leydesdorff, L., 2005. Network structure, self-organization, and the growth of international collaboration in science. *Res. Policy* 34, 1608–1618. doi:10.1016/j.respol.2005.08.002

CHAPTER 9

Careers, Labour Market and Individual Performance



Information sources – information targets: evaluative aspects of the scientists' publication strategies

Wolfgang Glänzel^{*,**}, Pei-Shan Chi^{*}, Christian Gumpenberger^{***} and Juan Gorraiz^{***}

^{*} wolfgang.glanzel@kuleuven.be; peishan.chi@kuleuven.be;
ECOOM and Dept MSI, KU Leuven, Naamssestraat 61, Leuven 3000, (Belgium)

^{**} glanzw@iif.hu;
Library of the Hungarian Academy of Sciences, Dept. Science Policy and Scientometrics, Arany János u. 1,
Budapest, 1051 (Hungary)

^{***} christian.gumpenberger@univie.ac.at; juan.gorraiz@univie.ac.at;
University of Vienna, Vienna University Library, Bibliometrics and Publication Strategies, Boltzmanngasse 5,
Vienna, 1090 (Austria)

ABSTRACT

Journal citation measures, if properly used, provide important information on the author's publication strategy. In this explorative study, which is part of a larger project, we attempt to shed light on to what extent publication strategies are adequately reflected by the impact generated in the respective scientific community in the context of academic research assessment at micro level. In this paper we present three cases based on the research output of researchers active in three different fields: chemistry, medicine and economics. In each individual case, the lists of journals, in which the author in question has published along with the journals in the reference lists and those where the citing papers have been published, are analysed according to two aspects, the congruence of the three resulting lists and the overlap by journal quartiles based on field-normalised impact. Similarity measures are then introduced at both levels. The results reveal important aspects of the authors' publication strategy and their position in the information flow enabling the identification of different scenarios, which are discussed in detail in order to be correctly applied for bibliometric individual assessment.

INTRODUCTION

Assessment of research performance at the micro level, notably of research teams and individual scientists has become increasingly important. Along with the widely accessible bibliographic or pre-processed data and readymade bibliometric tools, the application of indicators to the evaluation of individuals and their performance has become a compelling, assumedly feasible but basically challenging endeavour. The available web-based and sometimes poorly documented tools of variable quality and lacking standards have not yet proved serious alternatives for general purpose. The main issues in the context of indicator use, in general, and the use of readymade products, in particular, for individual level evaluation have been discussed, among by Wouters et al. (2013).

The discussion pointed to the limitations but also to the potential and opportunities of research assessment at this level: The utmost caution and accuracy of data and methods is required and bibliometrics has always to be applied in combination with subjective, peer-review based methods. In the bibliometric component of assessment the use of however sophisticated indicators is never sufficient. Professional profiles and research environment including co-

author networks are important aspects in the evaluation process (Wouters et al., 2013; Glänzel, 2014).

Other important aspects are the purpose and the target group. In practice it does matter and does make a difference if assessment is made, e.g., for possible promotion, for funding, for bestowing an award or in the context of a benchmarking exercise. Also the target group matters, firstly the scientists themselves, this is especially relevant for young scientists to inform them about the measurable aspects of their standing in the community and their publication strategies and how to improve them, and secondly the peers involved and selected for the evaluation of their research output, in order to help them providing a solid and reliable quantitative description of the scientist's research output and enabling to focus with their expertise on the qualitative aspects.

Publication activity and citation impact are the most favourite tools but some of the indicators, such as the h-index and the journal impact factor are because of frequent inappropriate use under dispute, and justifiably so. Yet, journal citation measures, if properly used, provide important information on the author's publication strategy. Another aspect of publication strategy is revealed by co-authorship patterns (cf. Schubert, 2012; Schubert, 2013; Glänzel, 2014), one important reason why Hirsch (2010) has modified his h-index.

At some universities, like at the University of Vienna, individual assessment also includes a visibility analysis of the scientist or the corresponding publications, which needs to be clearly distinguished from impact analyses based on the citations attracted by the scientists' publications.

Visibility analyses are helpful, whenever assessment exercises are performed for the last, most recent years, and the citation window is too short for retrieving a significant number of citations in many disciplines. This is particularly true for fields with a long cited half-life.

A visibility analysis comprises of three parts: first, the number and percentage of publications indexed in the different international, well-respected selected data sources; second, the number and percentage of publications in top journals or sources; and third, the number and percentage of publications in Open Access sources (Gorraiz and Gumpenberger, 2015; Gorraiz, Wieland and Gumpenberger, 2016)¹.

In this study we focus on the second aspect of the beforehand described approach, relying on the hypothesis that the visibility of a document is determined by the reputation or the impact of the source where it was published. By reflecting this editorial barrier, publication strategies can be unveiled.

The identification of the top journals is normally based on journal impact measures, like Garfield's impact factor² (Garfield, 1972 & 2005; Glänzel and Moed, 2002).

Therefore, the journals or sources used by the researcher under evaluation as publication channels are analysed and compared for different time periods. In a complementary reference analysis, the most cited sources and journals are determined, analysed (percentage of top journals) and compared with the previously identified publication channels (see visibility analysis). A good match is a strong indication that the scientist under evaluation has been successful to publish in the most relevant sources of his research area.

¹ Promotion strategies related to "altmetrics" are considered separately.

² Of course, other journal impact measures like "Article Influence Score", "SJR" or "SNIP" can also be used, depending on the data source (Scopus or Web of Science Core Collection).

Finally, the citing journals are retrieved as well and matched with the results of the previous analyses.

The main purpose of this explorative analysis is to promote a possible approach of how to quantify and interpret such data in a responsible way.

Moreover the question arises to what extent visibility is related to impact, or put differently, if publication strategies are adequately reflected by the impact generated in the respective scientific community. We attempt to shed light on this question in the context of academic research assessment at micro level.

METHODS

More particularly, we try to analyse individual strategies in the context of information flow, that is, in the context of information “sources” and “targets” (cf. Glänzel and Schoepflin, 1999). We assume a simple model, namely we consider the references cited by an author’s information sources, which might have been published in prominent journals or in contrast in less visible or even obscure ones. This will be contrasted by the publication venues the author has chosen for the own publication. The same will be done for the information targets, i.e., the citing journals. We just mention in passing that authors in several fields, e.g., in mathematics, aim at submitting their articles to that same journal in which the main information source was published – a strategy not always crowned with success.

In order to quantify data, we have applied two simple but robust measures. We focus only on publications indexed in and journals covered by Thomson Reuters’ Web of Science Core Collection (WoS). We list all journals in which a selected author has published along with the journals in the reference lists and where the citing papers have been published. Journals are weighted with the number of publications and then the three lists are merged to identify the overlap. All references and citations from the publication year till present are taken into account. In addition, we have calculated the average impact factor over the last ten years to be able to compare the visibility of references, publications and citations (Gorraiz et al., 2012). In particular, we have assigned all journals to their quartiles according to this impact measure within the same discipline.

In order to assess the author’s publication strategy, two aspects were analysed. First the congruence of the lists, which means, if an author tends to publish in the same journals that he/she cites and if the authors is cited in the same journals where he/she publishes. The second measure refers to the overlap by quartiles. This is assumed to provide information on the congruence of visibility, which might in principle be independent and different from the first approach. To give a trivial example, an author cites *Science* but prefers to publish in *Nature*

We are illustrating this method using three examples from the University Vienna. We have chosen one author each active in *chemistry*, *medicine* and *economics*. In order to measure the overlap between the three lists we have used the ratio of the geometric mean of the number of papers in the journal in question and of the geometrics mean of all WoS-indexed journal papers in the corresponding lists. The measure has similar properties as Salton’s measure: it ranges between 0 and 1 and the sum over the total will not exceed the value 1, which corresponds to complete congruence. The present measure also stresses the weight of the corresponding category, that is, if the distribution of journals in the two lists is similar,

journals with higher frequency obtain higher values as well. Finally, this is what we are interested in. The results for are shown in Table 1. Column “Q” gives the journals’ quartile.

RESULTS

The authors in chemistry and medicine published in a large range of journals, with considerable overlap between cited and citing journals mostly in the upper quartile. The overlap between cited and citing journals is distinctly higher. Interestingly, the economist has more homogeneous patterns but in this case most frequently used journals belong to the second or even lower quartiles. The overall congruence with regard to papers-references, papers-citations and references-citations is practically identical.

In a second step we have applied the same procedure to journal quartiles. Individual journals are replaced by the quartiles according to their average impact factor. Note that (weighted) similarities have to be recalculated as they cannot directly be from the previous case. The results are shown in Table 2. The congruence of quartiles is very strong in all three cases but here, again, overlap is in favour of the first quartile for researcher #1 and #3 and the economist tends to be represented by publications, references and citations in quartiles that are spread over the first three quartiles. Thus all three authors tend to publish in journals of the same levels as those they cite and by those their papers are cited.

Table 1. Distribution of overlap and weighted similarity by journals: chemistry (top), economy (centre) and medicine (bottom)

PUB-CITED overlap				PUB-CITING overlap				CITED-CITING overlap			
Papers	Cited	Q	Similarity	Papers	Citing	Q	Similarity	Cited	Citing	Q	Similarity
3	76	1	0.064	3	19	1	0.049	94	20	1	0.047
3	41	1	0.047	3	15	1	0.043	76	19	1	0.042
1	94	1	0.041	3	13	2	0.040	79	11	2	0.032
1	79	1	0.038	3	12	1	0.039	65	8	1	0.025
3	20	1	0.033	3	9	2	0.034	79	5	1	0.022
3	19	1	0.032	3	7	1	0.030	41	7	1	0.019
3	18	1	0.031	1	20	1	0.029	19	15	1	0.018
3	10	2	0.023	1	17	1	0.027	19	14	1	0.018
1	19	1	0.018	1	14	1	0.024	18	12	1	0.016
1	19	2	0.018	2	7	2	0.024	19	11	2	0.016
2	6	2	0.015	3	4	1	0.022	8	17	1	0.013
1	10	1	0.013	1	11	2	0.021	32	3	1	0.011
3	3	2	0.013	2	4	1	0.018	10	9	2	0.010
1	8	1	0.012	1	6	1	0.016	20	4	1	0.010
1	7	1	0.011	1	6	3	0.016	11	6	2	0.009
1	6	2	0.010	1	5	1	0.014	10	6	1	0.008
2	3	1	0.010	1	4	1	0.013	20	3	2	0.008
1	4	1	0.008	1	3	1	0.011	44	1	1	0.007
1	2	3	0.006	1	2	2	0.009	6	7	2	0.007
1	2	1	0.006	1	2	2	0.009	3	13	2	0.007
1	2	2	0.006	1	1	1	0.006	13	3	1	0.007
1	2	1	0.006	1	1	1	0.006	10	3	1	0.006
1	1	1	0.004	—	—	—	—	<0,006
Overall			0.467				0.501				0.574

PUB-CITED overlap				PUB-CITING overlap				CITED-CITING overlap			
Papers	Cited	Q	Similarity	Papers	Citing	Q	Similarity	Cited	Citing	Q	Similarity
4	30	2	0.085	3	16	2	0.074	76	6	1	0.056
1	76	1	0.068	4	11	2	0.070	21	16	2	0.048
4	16	4	0.062	2	12	2	0.052	30	11	2	0.048
3	21	2	0.062	4	4	4	0.042	22	13	2	0.045
2	11	2	0.037	2	8	4	0.042	30	9	1	0.043
1	22	2	0.037	1	13	2	0.038	11	12	2	0.030
1	14	1	0.029	1	11	2	0.035	14	6	1	0.024
1	13	1	0.028	1	6	1	0.026	11	6	1	0.021
1	12	1	0.027	1	6	1	0.026	16	4	4	0.021
1	11	1	0.026	1	6	1	0.026	5	11	2	0.020
1	9	1	0.023	1	5	1	0.024	13	4	1	0.019
1	7	1	0.021	1	4	1	0.021	12	2	1	0.013
1	5	2	0.017	1	4	3	0.021	4	4	3	0.011
1	4	4	0.016	1	2	2	0.015	7	2	1	0.010
1	4	3	0.016	1	2	1	0.015	1	13	2	0.010
1	2	2	0.011	1	2	2	0.015	6	2	3	0.009
1	2	1	0.011	1	2	1	0.015	6	2	1	0.009
1	1	4	0.008	1	2	4	0.015	4	3	3	0.009
—	—	—	—	1	1	4	0.011	10	1	1	0.008
—	—	—	—	1	1	3	0.011	<0,008

PUB-CITED overlap				PUB-CITING overlap				CITED-CITING overlap			
Papers	Cited	Q	Similarity	Papers	Citing	Q	Similarity	Cited	Citing	Q	Similarity
5	124	1	0.090	5	25	1	0.052	124	25	1	0.056
8	52	1	0.074	2	35	1	0.039	42	35	1	0.039
4	27	2	0.038	4	17	2	0.038	51	13	1	0.026
2	42	1	0.033	8	6	1	0.032	27	17	2	0.022
4	11	1	0.024	2	12	2	0.023	52	6	1	0.018
2	19	1	0.022	2	11	1	0.022	20	14	1	0.017
2	15	1	0.020	4	5	1	0.021	19	11	1	0.015
1	30	1	0.020	1	14	1	0.017	30	6	1	0.013
2	14	2	0.019	2	6	2	0.016	23	7	1	0.013
1	20	1	0.016	2	5	2	0.015	19	7	1	0.012
1	16	1	0.014	1	9	1	0.014	16	8	1	0.011
2	8	2	0.014	1	8	1	0.013	23	5	1	0.011
1	10	1	0.011	2	4	3	0.013	27	4	1	0.010
1	8	1	0.010	1	6	1	0.011	11	7	1	0.009
2	3	2	0.009	1	4	1	0.009	15	5	2	0.009
2	3	2	0.009	2	2	4	0.009	24	3	1	0.009
2	3	3	0.009	1	3	1	0.008	24	3	1	0.009
1	5	1	0.008	1	3	1	0.008	14	5	2	0.008
1	5	1	0.008	1	3	2	0.008	23	3	1	0.008
1	3	1	0.006	1	3	1	0.008	3	22	1	0.008
1	3	1	0.006	2	1	2	0.007	5	11	1	0.007
1	2	2	0.005	1	2	1	0.007	11	5	1	0.007
1	2	1	0.005	1	2	2	0.007	6	9	1	0.007
2	1	4	0.005	1	2	3	0.007	13	4	1	0.007
1	2	1	0.005	1	2	1	0.007	10	5	2	0.007
1	2	1	0.005	2	1	1	0.007	8	6	2	0.007
1	1	2	0.004	1	1	1	0.005	9	5	1	0.007
1	1	1	0.004	—	—	—	—	5	9	1	0.007
1	1	3	0.004	—	—	—	—	3	12	2	0.006
1	1	2	0.004	—	—	—	—	<0,006
Overall			0.502				0.420				0.616

Table 2. Distribution of overlap and weighted similarity by quartiles: chemistry (top), economy (centre) and medicine (bottom)

Discipline	Quartile	PUB-CITED	PUB-CITING	CITED-CITING
Chemistry	Q1	0.726	0.665	0.715
	Q2	0.224	0.261	0.212
	Q3	0.040	0.059	0.048
	Q4	0.000	0.000	0.010
	Overall	0.991	0.985	0.986
Economy	Q1	0.397	0.308	0.488
	Q2	0.314	0.392	0.303
	Q3	0.087	0.104	0.097
	Q4	0.108	0.178	0.077
	Overall	0.906	0.982	0.965
Medicine	Q1	0.722	0.630	0.699
	Q2	0.189	0.247	0.186
	Q3	0.057	0.079	0.067
	Q4	0.016	0.042	0.020
	Overall	0.983	0.997	0.973

DISCUSSION AND FUTURE RESEARCH

The “pub-cited” similarity (pcds) informs to what extent researchers are successful to publish in top journals in their research field(s). Now this definition of top journals does no longer only rely on bibliometric calculations, like Garfield's Impact Factor, but has been expanded by the information provided by the researcher under evaluation himself. In our approach, the journals most often cited by the researchers themselves are consequently considered as the most relevant sources in their specific research field(s). In so doing, potential errors or limitations of traditional journal impact measures, like incorrect category assignment or incomplete field delineation, can be mitigated.

Researchers are expected to successfully publish to a great extent in the top journals, which they also cite regularly and which constitute their knowledge base.

The “pub-citing” similarity (pcgs) gives an idea about the prestige of the citing journals in comparison to the journals where the researcher has published in.

Finally, the “cited-citing” similarity (ccs) provides insight to what extent the researcher's compiled knowledge base is in agreement with the resulting impact and the created knowledge transfer in the corresponding research field(s)³.

Using quartiles, and especially quartile one (Q1), allows a quantification of this information. Accordingly four different scenarios can be distinguished:

³ The latter data are only relevant if the citation window is big enough in comparison to the cited half-life of the research field.

- a) High percentage of publications in Q1 journals, high percentage of cited references in Q1 journals, and high percentage of attracted citations in Q1 journals.
This scenario hints at a top performer and confirms a sound publication strategy. The more the similarity values exceed the expected ones (0.25), the higher is the expected performance and the more successful is the present publication strategy.
- b) Low percentage of publications in Q1 journals, high percentage of cited references in Q1 journals, and low percentage of attracted citations in Q1 journals.
Similarity values below 0.25 hints at a low performance reflected in the present publication strategy.
- c) Low percentage of publications in Q1 journals, high percentage of cited references in Q1 journals, and high percentage of attracted citations in Q1 journals.
This scenario suggests that the publication strategy could be improved, which would most probably also positively influence the performance.
- d) Low percentage of publications in Q1 journals, low percentage of cited references in Q1 journals, and low percentage of attracted citations in Q1 journals.
This scenario does not automatically mean a low performance. It is rather advisable to invest time in a thorough delineation of the research field as well as in critically reviewing the currently defined top journals (Q1).

In the three case studies presented in this paper, the researchers in chemistry and medicine are good examples for the first scenario with high values of all three similarity values. In contrast, the researcher in economics, despite slightly exceeding the similarity threshold in Q1, seems to belong to a different league. This is reflected in the similarity values for the second quartile.

These results are in good agreement with our citation analyses performed a posteriori, which show a much higher category normalized citation impact and a higher number and percentage of publications in the Top 1% and Top 10% most cited papers for the researchers in chemistry and medicine than for the researcher in economics.

It should be stressed that visibility analyses are only used in order to provide a quantitative description of the research output and to reveal potentially meaningful symptoms. Of course, researchers might always have good reasons for their choices of publication channels. However, particularly junior scientists should be made aware of the consequences of careless or even ill-conceived publication strategies.

Forthcoming studies will include more researchers from further disciplines. This is necessary for further clarification and validation of the obtained results.

Furthermore, there are two important aspects we will include in future evaluative tasks. The first issue was already mentioned above, the analysis of the position of the authors in their co-author networks using the models proposed by Glänzel (2014) and Schubert (“molecular model”; 2013). The second aspect is based on a new method based on the already established methods of Characteristic Scores and Scales (Glänzel and Schubert, 1988) that proved to be applicable to the micro level as well (Thijs et al., 2014). This combination promises to

provide a bibliometric portfolio to be used to support decision making in combination with peer review and expert opinion.

REFERENCES

- Garfield, E. (2005). The Agony and the Ecstasy. The History and Meaning of the Journal Impact Factor. <http://garfield.library.upenn.edu/papers/jifchicago2005.pdf>; [Accessed on March 2016]
- Garfield, E. (1972). Citation analysis as a tool in journal evaluation: Journals can be ranked by frequency and impact of citations for science policy studies', *Science*, 178(4060), 471–479.
- Glänzel, W. & Schubert, A. (1988), Characteristic Scores and Scales in assessing citation impact. *Journal of Information Science*, 14(2), 123–127.
- Glänzel, W., Schoepflin, U. (1999), A bibliometric study of reference literature in the sciences and social sciences. *Information Processing and Management*, 35(1), 31–44.
- Glänzel, W. (2014), On co-authorship at the individual level. *Transinformação* , 26(3), 229–238.
- Glänzel, W., & Moed, H. F. (2002). Journal impact measures in bibliometric research. *Scientometrics*, 53(2), 171–193.
- Gorraiz, J., Gumpenberger, C., Schlögl, C., & Wieland, M. (2012). On the temporal stability of Garfield's Impact Factor and its suitability to identify hot papers. *Proceedings of STI 2012 Montreal. 17th International Conference on Science and Technology Indicators*, 1, 319–332.
- Gorraiz, J. and Gumpenberger, C. (2015). A flexible bibliometric approach for the assessment of professorial appointments. *Scientometrics*, 105, 3, 1699-1719. DOI 10.1007/s11192-015-1703-6
- Gorraiz, J. and Gumpenberger, C. (2016). Individual Bibliometric Assessment @ University of Vienna: From Numbers to Multidimensional Profiles; <http://arxiv.org/abs/1601.08049>, [Accessed on March 2016]
- Schubert, A. (2012), A Hirsch-type index of co-author partnership ability. *Scientometrics*, 91(1), 303–308.
- Schubert, A. (2013), *Atoms and molecules of scientometrics. Authors, publications, references, citations and the bonds among them*. Toulouse, 2013.
- Thijs, B., Debackere, K., Glänzel, W (2014). Improved author profiling through the use of citation classes. In: E. Noyons (Ed.), *Proceedings of the STI Conference 2014*, Leiden, 616–622
- Wouters, P., Glänzel, W., Gläser, J., Rafols, I. (2013), The Dilemmas of Performance Indicators of Individual Researchers – An Urgent Debate in Bibliometrics. *ISSI Newsletter*, 9 (3), 48–53.



The Effect of Holding a Research Chair on Scientists' Research Impact

Seyed Reza Mirnezami^{*}, and Catherine Beaudry^{**}

^{*}seyed-reza.mirnezami@polymtl.ca
Polytechnique Montreal, Canada

^{**}catherine.beaudry@polymtl.ca
Polytechnique Montreal, Canada

ABSTRACT

This paper examines the effect of holding Canada Research Chair (CRC) on a scientist's number of citations as a measure of research impact, based on an econometric analysis with combined data on Quebec scientists' funding and journal publication. Using Generalized Least Square (GLS) method for regression analysis, the results show that holding either tier-1 or tier-2 of CRC significantly and positively results in conducting research with higher impact. This finding, however, does not necessarily imply that the others are the lesser scientists.

Keywords: Citation, Research Chair, Research Impact, Science Policy.

INTRODUCTION

Research impact is an important topic in science policy. The administrative bodies and policy makers want to get the maximum benefit of public budget, which tax-payers spend for the sake of knowledge production and contribution to the advancement of knowledge. Considering the standing of research impact in science policy issues, it is fruitful to investigate the determinants of citation count. Various factors have been mentioned in the literature that explain the number of citation: The size of research team or measure of research collaboration (Johnes, 1988; Melin, 1996), the research domain, the prestige of the journal, and the social network of authors (Bornmann et al., 2008), the scientist gender (Aksnes et al., 2011), the amount of research funding (Harman, 2000; Pavitt, 2000, 2001), and scientist visibility in academic community (Mirnezami et al., 2015).

Conducting an econometric analysis, this paper identifies the main determinant of citation count, specifically looking at the effect of 'holding a research chair' on citation count. The remainder of this article goes as follows: Section 2 reviews the related blocks in the literature; Section 3 introduces the data set and explains the research methodology; Section 4 presents the regression analyses; and finally, Section 5 discusses the results and concludes findings.

SECTION 1 - THEORETICAL FRAMEWORK

In order to situate the topic of this research, we review two related blocks of literature: 'prestigious academic affiliation' and 'the number of citation' as a measure of research impact.

An explanation for the covariation of ‘research quality’ and ‘the number of citation’ has been provided by Kostoff (1998) and similarly by Phelan (1999). Kostoff (1998) investigates the theory of citation and proposes that each citation has more or less two origins/components: the real component of intellectual heritage and random components of self-interest. The author argues that the random effect diminishes in the aggregation of citation counts and therefore the number of citation is a good indicator of the “research quality”.

In terms of prestige, Long et al. (1979) showed a positive and significant correlation between the prestige of the scientist alma matter/affiliation and the number of citations. Honors and awards can be also proxies for research prestige, if they are given/awarded based on competitive and pre-defined procedures, like what is called as ‘research chair program’ in Canada. Cantu et al. (2009) showed the research chair programs are capable of implementing knowledge-based development. Considering holding a chair as a measure of prestige, we examine the effect of being a ‘chair-holder’ on research impact. Our hypothesis therefore reads as:

Hypothesis: *Holding a chair increases a scientist’s research impact measured in terms of number of citations.*

There are some other factors mentioned in the literature as possible determinants of research impact. These can be used as control variables in regression analysis. The age of scientist may affect the scientific productivity (Kyvik, 1990; Kyvik and Olsen, 2008). Gender is also known as a significant determinant of scientific productivity in the literature (Long, 1990, 1992). Research funding can be another determinant (Salter and Martin, 2001).

Other factor which have been mentioned in the literature are the size of research team and department size (Buchmueller et al., 1999; Carayol and Matt, 2006; Heinze et al., 2009), the type of university governance and ownership (Golden and Carstensen, 1992; Jordan et al., 1989), attributions of each specific research field and scientific context, which may characterize the research impact (Baird, 1991; Blackburn et al., 1978), and scientist visibility in form of number of articles or average impact factor of journals in which scientists publish his/her articles (Feist, 1997; Merton, 1968; Stegmann and Grohmann, 2001).

SECTION 2 - DATA AND METHODOLOGY

Methodology and econometrics model

Our regression analysis aims to explain the number of citation as the left-hand-side (LHS) variable based on the right-hand-side (RHS) variables, which are reviewed above. To measure the effect of ‘holding a research chair’ on a scientific research impact/quality, we use Generalized Least Squares (GLS) model. This is a technique for linear regression models, used when there is a certain degree of correlation between the residuals in a regression model. In other words, the variance matrix of dependent variable is no longer a scalar variance-covariance matrix. The following graph in Figure 1 shows that the standard deviation of citation count is not constant over ages. In such circumstances, OLS and WLS are statistically inefficient, which give misleading inferences. The command of *xtgls* in STATA fits GLS models on the panel data.

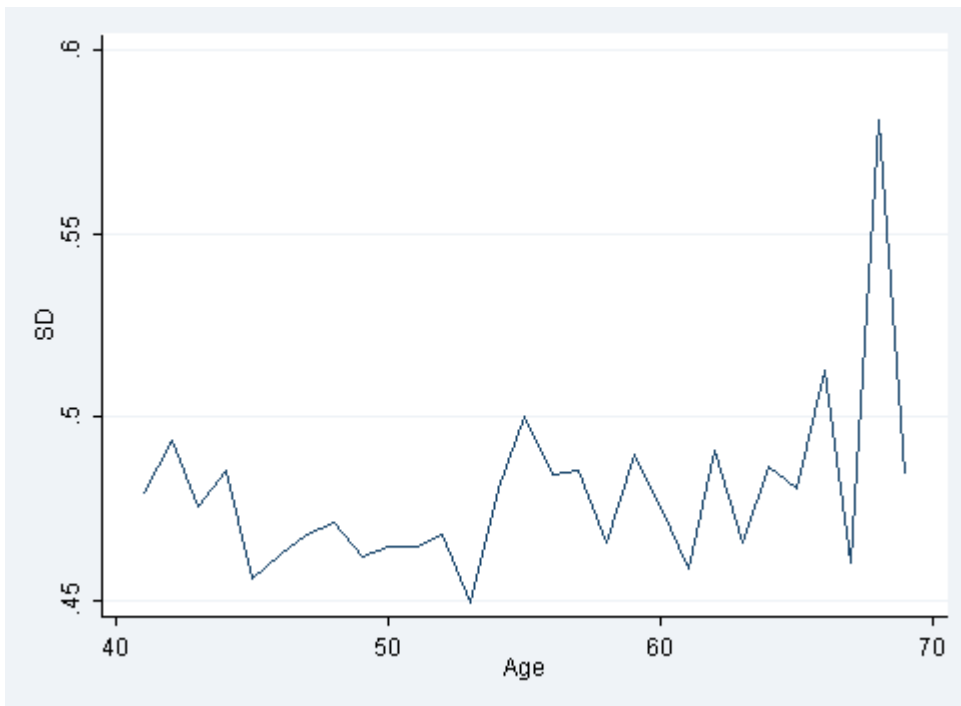


Figure 1 - The variation of citation count standard deviation over ages

In addition to CRC as the main independent variable, we also put some control variables in the model: the amount of funding as the scientists' operational capacity to conduct research, the number of articles and journal impact factor as measures of scientists past performance/visibility and his/her experience, the average number of authors in articles indicating the size of academic network, and the average of citation count for the first three years reflecting the initial condition of researcher. The use of initial condition to improve model efficiency has been verified in Blundell and Smith (1990). In addition, we put the gender of scientist in the left-hand-side to control some un-observed characteristics of them. Finally, dummy variables of universities and years are also put in the model to consider institutional effect on scientists' performance. Figure 2 shows the different average of scientists in universities justifying use of dummy variables in our model. Figure 3 justifies the use of year dummies in our model.

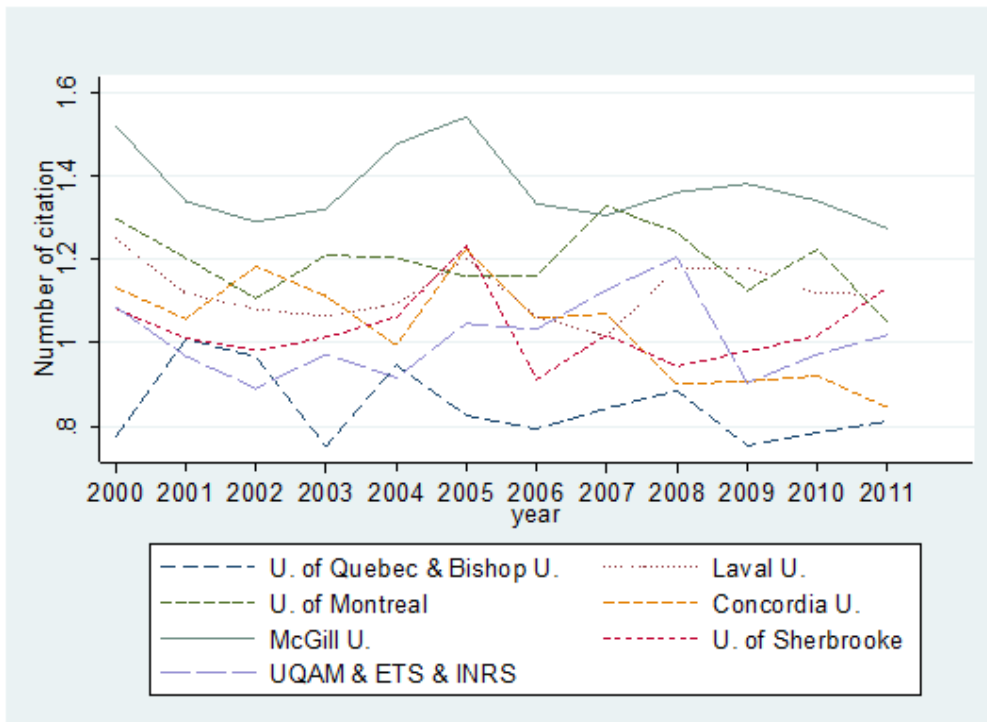


Figure 2 - Discipline-normalized citation rates of papers from Quebec universities



Figure 3 - Average of Discipline-normalized citation rates of Quebec papers, by year

Considering the mentioned explanatory variables, the resulting model is given by

$$\ln(\text{nbCitation}) = f \left(\begin{array}{l} \text{Tier 1, Tier 2, Citavg 3, } \ln(\text{PublicfundingO}), \ln(\text{PublicfundingI}) \\ \ln(\text{PrivatefundingO}), \ln(\text{NFPfundingO}), \ln(\text{nbArticle}), \\ \ln(\text{nbAuthor}), \ln(\text{Impactfactor}), d\text{Female}, \text{Age}, d\text{Univ} \end{array} \right)$$

Data and variables

The data set used in this article integrates information about funding and publication of scientists in the province of Quebec. Funding information of scientists comes from the Quebec University Research Information System (*Système d'information sur la recherche universitaire* or SIRU) of the Ministry of Education, Leisure and Sports (MELS). This database reports funding information including research grants and industrial contracts of all Quebec academics, on a yearly basis during the period 1985-2012. We have access to Thomson Reuters Web of Science database on scientific articles (2000-2012), which includes information about date of publication, journal name, authors, affiliations, and the number of citation each article receives. To identify chair holders, we got information of all chair holders from Canada Research Chair office¹.

SECTION 3 - RESULT AND DISCUSSION:

The result of regression analysis in Table 1, Table 2, and Table 3 show that, *ceteris paribus*, both tiers of CRC have significant and positive effect on research impact. No matter which tier of CRC a scientist has, such chair holding is a kind of proxy for latent variables indicating the inherent capabilities in conducting research. To justify this finding, we can argue that CRC is a prestigious research sign in Canada, which grants more visibility to the chair-holders. As a result, chair holders are almost successful in academic networking and attracting accomplished and promising minds in academia. In addition, non-chair holders may also have more willingness to conduct collaborative research with the CRC holders as they have well-equipped laboratories and talented research staff. Regarding the effect of initial condition, the results show that the average of citation counts for the first three years positively affect the number of citation in future. As mentioned in previous section, it is a technique to increase the efficiency of our dynamic panel model.

Beside the effect of chair holding, there are some significant effect of control variables. The variable of [*dFemale*] is significant with a negative effect on the number of citations. However, when we consider the interactive effect of gender with the amount of public funding and with the number of articles, the results suggest that female with low amount of funding or few articles are being cited more than male while other female scientists are cited less than male. This finding can be related to Aksnes et al. (2011) and Larivière et al (2013) showing the underperformance of women.

¹ Tier 1 and Tier 2 CRC holders receive annual amount of \$200,000 and \$100,000 respectively.

Table 1 – GLS regression results to investigate the effect of tier-1 CRC on citation count

<i>Dep var: ln(nbCitation)_{it}</i>	<i>Reg1</i>	<i>Reg2</i>	<i>Reg3</i>	<i>Reg4</i>	<i>Reg5</i>	<i>Reg6</i>	<i>Reg7</i>
<i>Citavg3</i>	0.0903*** 0.0020	0.0854*** 0.0025	0.0888*** 0.0021	0.0921*** 0.0020	0.0868*** 0.0023	0.0902*** 0.0021	0.0902*** 0.0021
<i>Tier1_{it}</i>	0.0528*** 0.0085	0.0498*** 0.0084	0.0495*** 0.0084	0.0538*** 0.0085	0.0511*** 0.0085	0.0385 0.0580	0.0382*** 0.0119
<i>Tier1_{it}*ln(PublicfundingO)_{it}</i>						0.0012 0.0047	
<i>Tier1_{it}*ln(PrivatefundingO)_{it}</i>							0.0028* 0.0016
<i>dFemale</i>	-0.0098*** 0.0017	0.0167*** 0.0048	0.0224*** 0.0078	-0.0102*** 0.0018	-0.0068*** 0.0020	-0.0096*** 0.0017	-0.0096*** 0.0017
<i>ln(PublicfundingO)_{it}</i>	-0.0014*** 0.0003	-0.0020*** 0.0002	-0.0013*** 0.0002	-0.0014*** 0.0003	-0.0016*** 0.0003	-0.0015*** 0.0003	-0.0015*** 0.0003
<i>ln(PrivatefundingO)_{it}</i>	-0.0001 0.0002	0.0001 0.0002	-0.0005** 0.0002	-0.0002 0.0003	0.0002 0.0003	-0.0001 0.0002	-0.0002 0.0002
<i>ln(NFPfundingO)_{it}</i>	-0.0013*** 0.0002	-0.0015*** 0.0002	-0.0010*** 0.0002	-0.0013*** 0.0002	-0.0011*** 0.0002	-0.0013*** 0.0002	-0.0013*** 0.0002
<i>ln(nbArticle)_{it}</i>	0.0643*** 0.0020	0.0750*** 0.0024	0.0676*** 0.0020	0.0637*** 0.0020	0.0564*** 0.0023	0.0644*** 0.0020	0.0643*** 0.0020
<i>ln(Impactfactor)_{it}</i>	0.3530*** 0.0013	0.3495*** 0.0013	0.3534*** 0.0013	0.3525*** 0.0014	0.3290*** 0.0019	0.3528*** 0.0013	0.3528*** 0.0013
<i>ln(nbAuthor)_{it}</i>	0.1080*** 0.0012	0.1115*** 0.0015	0.1071*** 0.0012	0.1081*** 0.0014	0.1060*** 0.0016	0.1079*** 0.0012	0.1079*** 0.0012
<i>dFemale*ln(nbArticle)_{it}</i>		-0.0214*** 0.0043					
<i>dFemale*ln(PublicfundingO)_{it}</i>			-0.0033*** 0.0008				
<i>dFemale*ln(PrivatefundingO)_{it}</i>				0.0005 0.0006			
<i>ln(Impactfactor)_{it}*ln(nbArticle)_{it}</i>					0.0126*** 0.0010		
<i>Constant</i>	0.3135*** 0.0039	0.3031*** 0.0045	0.3121*** 0.0036	0.3109*** 0.0042	0.3156*** 0.0044	0.3131*** 0.0039	0.3134*** 0.0039
<i>Number of observation</i>	35332	35332	35332	35332	35332	35332	35332
<i>Number of scientists</i>	7315	7315	7315	7315	7315	7315	7315
χ^2	3345076	1033287	2978141	2452072	635693	3757878	2376291

Notes: *, **, and *** show the significance level at 0.1, 0.05, and 0.01 respectively - Year dummies, and university dummies are significant. The minimum year activity, average year activity, and maximum year activity are 1, 4.8, and 9 respectively.

Table 2 – GLS regression results to investigate the effect of tier-2 CRC on citation count

<i>Dep var: ln(nbCitation)_{it}</i>	<i>Reg8</i>	<i>Reg9</i>	<i>Reg10</i>	<i>Reg11</i>	<i>Reg12</i>	<i>Reg13</i>	<i>Reg14</i>
<i>Citavg3</i>	0.0896*** 0.0022	0.0887*** 0.0026	0.0904*** 0.0023	0.0915*** 0.0021	0.0893*** 0.0024	0.0903*** 0.0022	0.0899*** 0.0023
<i>Tier2_{it}</i>	0.0493*** 0.0090	0.0523*** 0.0090	0.0482*** 0.0090	0.0508*** 0.0090	0.0479*** 0.0091	0.1643*** 0.0588	0.0528*** 0.0111
<i>Tier2_{it}*ln(PublicfundingO)_{it}</i>						-0.0102** 0.0051	
<i>Tier2_{it}*ln(PrivatefundingO)_{it}</i>							-0.0015 0.0020
<i>dFemale</i>	-0.0082*** 0.0019	0.0211*** 0.0049	0.0230*** 0.0079	-0.0077*** 0.0020	-0.0069*** 0.0020	-0.0084*** 0.0019	-0.0080*** 0.0019
<i>ln(PublicfundingO)_{it}</i>	-0.0017*** 0.0003	-0.0024*** 0.0002	-0.0014*** 0.0003	-0.0018*** 0.0003	-0.0015*** 0.0003	-0.0017*** 0.0003	-0.0018*** 0.0003
<i>ln(PrivatefundingO)_{it}</i>	0.0000 0.0002	0.0004** 0.0002	-0.0004* 0.0002	0.0000 0.0003	0.0005** 0.0002	-0.0001 0.0002	0.0000 0.0002
<i>ln(NFPfundingO)_{it}</i>	-0.0013*** 0.0002	-0.0014*** 0.0002	-0.0009*** 0.0002	-0.0012*** 0.0002	-0.0013*** 0.0002	-0.0012*** 0.0002	-0.0012*** 0.0002
<i>ln(nbArticle)_{it}</i>	0.0701*** 0.0021	0.0770*** 0.0024	0.0716*** 0.0021	0.0690*** 0.0020	0.0600*** 0.0023	0.0704*** 0.0021	0.0701*** 0.0021
<i>ln(Impactfactor)_{it}</i>	0.3509*** 0.0014	0.3494*** 0.0013	0.3514*** 0.0014	0.3504*** 0.0015	0.3267*** 0.0020	0.3508*** 0.0014	0.3509*** 0.0014
<i>ln(nbAuthor)_{it}</i>	0.1066*** 0.0013	0.1098*** 0.0015	0.1065*** 0.0012	0.1064*** 0.0014	0.1068*** 0.0016	0.1063*** 0.0013	0.1063*** 0.0013
<i>dFemale*ln(nbArticle)_{it}</i>		-0.0264*** 0.0045					
<i>dFemale*ln(PublicfundingO)_{it}</i>			-0.0033*** 0.0008				
<i>dFemale*ln(PrivatefundingO)_{it}</i>				-0.0002 0.0006			
<i>ln(Impactfactor)_{it}*ln(nbArticle)_{it}</i>					0.0127*** 0.0010		
<i>Constant</i>	0.3092*** 0.0039	0.3022*** 0.0046	0.3052*** 0.0040	0.3083*** 0.0041	0.3048*** 0.0046	0.3091*** 0.0039	0.3093*** 0.0039
<i>Number of observation</i>	35332	35332	35332	35332	35332	35332	35332
<i>Number of scientists</i>	7315	7315	7315	7315	7315	7315	7315
<i>χ²</i>	2398663	430020	8692326	703639	415534	1877470	2307896

Notes: *, **, and *** show the significance level at 0.1, 0.05, and 0.01 respectively - Year dummies, and university dummies are significant. The minimum year activity, average year activity, and maximum year activity are 1, 4.8, and 9 respectively.

Table 3 – GLS regression results to investigate the effect of tier-1 and tier-2 CRC on citation count

<i>Dep var: ln(nbCitation)_{it}</i>	<i>Reg15</i>	<i>Reg16</i>	<i>Reg17</i>	<i>Reg18</i>	<i>Reg19</i>	<i>Reg20</i>	<i>Reg21</i>
<i>Citavg3</i>	0.0865*** 0.0023	0.0865*** 0.0026	0.0878*** 0.0024	0.0875*** 0.0022	0.0882*** 0.0025	0.0870*** 0.0023	0.0868*** 0.0023
<i>Tier1_{it}</i>	0.0576*** 0.0085	0.0541*** 0.0084	0.0549*** 0.0085	0.0579*** 0.0085	0.0544*** 0.0085	0.0451 0.0580	0.0445*** 0.0120
<i>Tier2_{it}</i>	0.0542*** 0.0090	0.0561*** 0.0090	0.0535*** 0.0090	0.0550*** 0.0090	0.0521*** 0.0091	0.1654*** 0.0589	0.0573*** 0.0112
<i>Tier1_{it}*ln(PublicfundingO)_{it}</i>						0.0010 0.0047	
<i>Tier1_{it}*ln(PrivatefundingO)_{it}</i>							0.0026* 0.0016
<i>Tier2_{it}*ln(PublicfundingO)_{it}</i>						-0.0099** 0.0051	
<i>Tier2_{it}*ln(PrivatefundingO)_{it}</i>							-0.0012 0.0020
<i>dFemale</i>	-0.0051*** 0.0019	0.0197*** 0.0050	0.0219*** 0.0079	-0.0044** 0.0020	-0.0049** 0.0021	-0.0054*** 0.0020	-0.0048** 0.0020
<i>ln(PublicfundingO)_{it}</i>	-0.0018*** 0.0003	-0.0022*** 0.0002	-0.0016*** 0.0003	-0.0019*** 0.0003	-0.0019*** 0.0003	-0.0018*** 0.0003	-0.0019*** 0.0003
<i>ln(PrivatefundingO)_{it}</i>	0.0001 0.0002	0.0007*** 0.0002	-0.0001 0.0002	0.0001 0.0003	0.0006** 0.0002	0.0000 0.0002	0.0001 0.0002
<i>ln(NFPfundingO)_{it}</i>	-0.0012*** 0.0002	-0.0015*** 0.0002	-0.0009*** 0.0002	-0.0012*** 0.0002	-0.0009*** 0.0002	-0.0011*** 0.0002	-0.0012*** 0.0002
<i>ln(nbArticle)_{it}</i>	0.0645*** 0.0022	0.0733*** 0.0024	0.0666*** 0.0022	0.0647*** 0.0022	0.0554*** 0.0024	0.0648*** 0.0022	0.0646*** 0.0022
<i>ln(Impactfactor)_{it}</i>	0.3519*** 0.0014	0.3484*** 0.0013	0.3512*** 0.0015	0.3511*** 0.0015	0.3285*** 0.0020	0.3514*** 0.0014	0.3513*** 0.0014
<i>ln(nbAuthor)_{it}</i>	0.1082*** 0.0013	0.1115*** 0.0015	0.1071*** 0.0012	0.1084*** 0.0015	0.1074*** 0.0016	0.1081*** 0.0014	0.1083*** 0.0013
<i>dFemale*ln(nbArticle)_{it}</i>		-0.0240*** 0.0045					
<i>dFemale*ln(PublicfundingO)_{it}</i>			-0.0029*** 0.0008				
<i>dFemale*ln(PrivatefundingO)_{it}</i>				-0.0004 0.0006			
<i>ln(Impactfactor)_{it}*ln(nbArticle)_{it}</i>					0.0123*** 0.0010		
<i>Constant</i>	0.3097*** 0.0040	0.3010*** 0.0048	0.3068*** 0.0041	0.3078*** 0.0042	0.3099*** 0.0047	0.3081*** 0.0041	0.3095*** 0.0041
<i>Number of observation</i>	35332	35332	35332	35332	35332	35332	35332
<i>Number of scientists</i>	7315	7315	7315	7315	7315	7315	7315
<i>χ²</i>	774094	307409	533071	1044830	239834	643384	595455

Notes: *, **, and *** show the significance level at 0.1, 0.05, and 0.01 respectively - Year dummies, and university dummies are significant. The minimum year activity, average year activity, and maximum year activity are 1, 4.8, and 9 respectively.

The number of articles [$\ln(nbArticle)$], journal impact factor [$\ln(Impactfactor)$], and interaction between them have all has a significant and positive effect on citation count. This implies that greater visibility of scientists can results in receiving more citations by them. It also shows that more articles in high impact factor journals results in more citations than the same number of articles in a less prestigious journal. Interestingly, the positive effect of visibility is smaller for female as shown in Figure 4. Related to our finding, there are some evidence in literature (Calderini and Franzoni, 2004; Stegmann and Grohmann, 2001) supporting the point that journal impact factor of past publication can be a proxy for research quality and visibility. However, one may criticize that journal impact factor is not a perfect proxy for research quality and research impact as citation count in journal has a significant variation with skewed distribution, which means that journal impact factor is based on few highly-cited items.

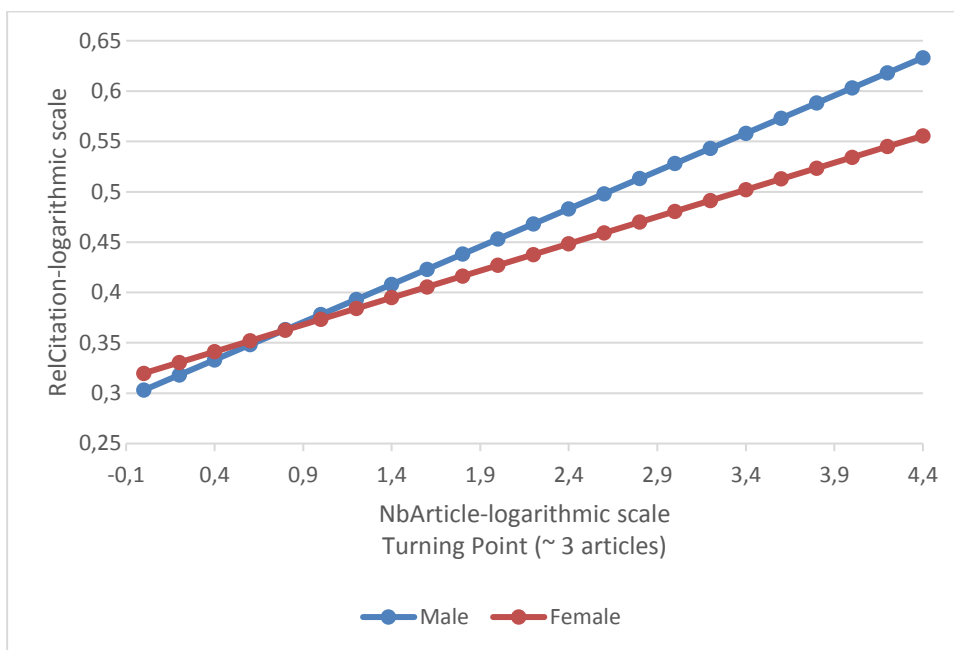


Figure 4 - Interactive effect of gender and number of articles

In terms of research team size, our results show that collaborative works with more authors [$\ln(nbAuthor)$] are more likely to be cited. The main reason for this finding is that collaborative nature of research work leads in higher quality, which is also supported by some articles (Johnes, 1988; Melin, 1996). This is mainly because tasks are broken down efficiently and research activities are being conducted in a collective way. On top of that, some sort of knowledge spillover or tacit knowledge transfer are possible by-product of such research collaboration, which improves their capability in conducting high impact research in future.

For the effect of funding, we got some mixed signals from our results. Although all of them show significant effect of funding on the citation count but only private funding [$\ln(PrivatefundingO)$] has some positive effect while funding form public sector [$\ln(PublicfundingO)$] or funding from non-profit organizations [$\ln(NFPfundingO)$] always have negative effect. The interactive effect of funding and gender is illustrated in Figure 5 - interactive effect of funding and gender Figure 5 showing that female has more negative effect of funding. Our previous empirical study on this database (Mirnezami and Beaudry, 2016) along with other evidences from literature (Arundel and Geuna, 2004; Harman, 2000; Pavitt, 2000, 2001), support the positive effect of funding on publication and scientific productivity of

scientists, but the results of this paper imply that higher funding does not necessarily results in publications which are more cited.

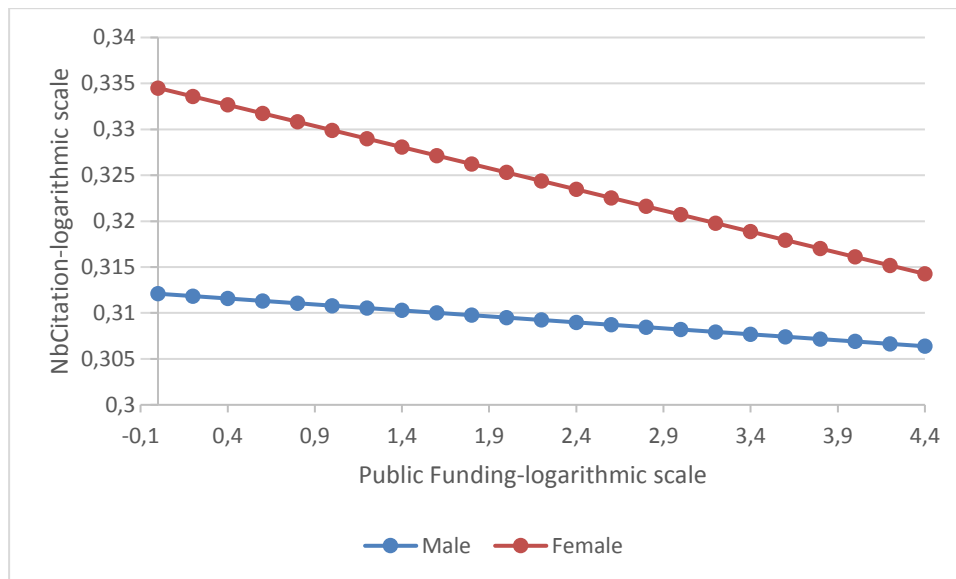


Figure 5 - interactive effect of funding and gender

CONCLUSION

The paper investigates the effect of holding CRC on citation count as a measure of research impact, which has been verified for both tier-1 and tier-2 of CRC. In addition, the positive effect of research team size, positive effect of number of articles and journal impact factor, and negative effect of being female have been validated based on our regression analysis. For funding effect, we both positive and negative effect depending on source of fund. We have also seen significant effect of year dummies and universities dummies, indicating the control of some un-observed institutional dimensions of research performance.

As a limitation to our mentioned interpretations, we only studied Quebec scientists and some data entries are missing in the original dataset. In addition to using more comprehensive and complete data set for future studies, one can conduct a deep investigation on citation concept and disentangle self-citation, citation based on quality, and citation related to research impact or literature review. In addition, future research can look for time-variation and discipline-dependency of our result or even investigate the effect of initial conditions on research impact/quality.

In terms of policy implication, we can conclude that CRC program is an effective strategy to improve research impact and the quality of research. In addition, one may argue that collaborative works (measured by the size of research team) should be encouraged in order to have scientific productivity with higher level of quality.

REFERENCES

- Aksnes, D.W., Rorstad, K., Piro, F., Sivertsen, G., 2011. Are female researchers less cited? A large-scale study of Norwegian scientists. *Journal of the American Society for Information Science and Technology* 62, 628-636.
- Arundel, A., Geuna, A., 2004. Proximity and the use of public science by innovative European firms. *Economics of Innovation and new Technology* 13, 559-580.

- Baird, L.L., 1991. Publication productivity in doctoral research departments: Interdisciplinary and intradisciplinary factors. *Research in Higher Education* 32, 303-318.
- Blackburn, R.T., Behymer, C.E., Hall, D.E., 1978. Research note: Correlates of faculty publications. *Sociology of Education*, 132-141.
- Blundell, R., Smith, R.J., 1990. Conditions initiales et estimation efficace dans les modèles dynamiques sur données de panel: Une application au comportement d'investissement des entreprises. *Annales d'économie et de Statistique*, 109-123.
- Bornmann, L., Mutz, R., Neuhaus, C., Daniel, H.-D., 2008. Citation counts for research evaluation: standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in Science and Environmental Politics* 8, 93-102.
- Buchmueller, T.C., Dominitz, J., Lee Hansen, W., 1999. Graduate training and the early career productivity of Ph.D. economists. *Economics of Education Review* 18, 65-77.
- Calderini, M., Franzoni, C., 2004. Is academic patenting detrimental to high quality research. An empirical analysis of the relationship between scientific careers and patent applications. Bocconi University: Cespri Working Paper.
- Cantu, F.J., Bustani, A., Molina, A., Moreira, H., 2009. A knowledge-based development model: the research chair strategy. *Journal of Knowledge Management* 13, 154-170.
- Carayol, N., Matt, M., 2006. Individual and collective determinants of academic scientists' productivity. *Information Economics and Policy* 18, 55-72.
- Feist, G.J., 1997. Quantity, quality, and depth of research as influences on scientific eminence: Is quantity most important? *Creativity Research Journal* 10, 325-335.
- Golden, J., Carstensen, F.V., 1992. Academic research productivity, department size and organization: Further results, comment. *Economics of Education Review* 11, 153-160.
- Harman, G., 2000. Allocating research infrastructure grants in post-binary higher education systems: British and Australian approaches. *Journal of Higher Education Policy and Management* 22, 111-126.
- Heinze, T., Shapira, P., Rogers, J.D., Senker, J.M., 2009. Organizational and institutional influences on creativity in scientific research. *Research Policy* 38, 610-623.
- Johnes, G., 1988. Determinants of research output in economics departments in British universities. *Research Policy* 17, 171-178.
- Jordan, J.M., Meador, M., Walters, S.J.K., 1989. Academic research productivity, department size and organization: Further results. *Economics of Education Review* 8, 345-352.
- Kostoff, R.N., 1998. The use and misuse of citation analysis in research evaluation. *Scientometrics* 43, 27-43.
- Kyvik, S., 1990. Age and scientific productivity. Differences between fields of learning. *Higher Education* 19, 37-55.
- Kyvik, S., Olsen, T.B., 2008. Does the aging of tenured academic staff affect the research performance of universities? *Scientometrics* 76, 439-455.
- Long, J.S., 1990. The origins of sex differences in science. *Social Forces* 68, 1297-1316.
- Long, J.S., 1992. Measures of sex differences in scientific productivity. *Social Forces* 71, 159-178.
- Long, J.S., Allison, P.D., McGinnis, R., 1979. Entrance into the academic career. *American Sociological Review*, 816-830.

- Melin, G., 1996. The networking university. *Scientometrics* 35, 15-31.
- Merton, R.K., 1968. The Matthew effect in science. *Science* 159, 56-63.
- Mirnezami, S.R., Beaudry, C., 2016. The effect of holding a research chair on scientists' productivity. *Scientometrics* 107, 399-454.
- Mirnezami, S.R., Beaudry, C., Larivière, V., 2015. What determines researchers' scientific impact? A case study of Quebec researchers. *Science and Public Policy*, scv038.
- Pavitt, K., 2000. Why European Union funding of academic research should be increased: a radical proposal. *Science and Public Policy* 27, 455-460.
- Pavitt, K., 2001. Public policies to support basic research: What can the rest of the world learn from US theory and practice?(And what they should not learn). *Industrial and corporate change* 10, 761-779.
- Salter, A.J., Martin, B.R., 2001. The economic benefits of publicly funded basic research: a critical review. *Research Policy* 30, 509-532.
- Stegmann, J., Grohmann, G., 2001. Citation rates, knowledge export and international visibility of dermatology journals listed and not listed in the Journal Citation Reports. *Scientometrics* 50, 483-502.

Table 4 - Variable description

Variable name	Variable description
<i>Tier1</i>	Dummy variables taking the value 1 if a scientist has a Canada research chair (tier 1)
<i>Tier2</i>	Dummy variables taking the value 1 if a scientist has a Canada research chair (tier 2)
<i>Citavg3</i>	The average of number of citations during the first three years
$\ln(nbCitation)_{it}$	Natural logarithm of number of citations of papers published by scientist <i>i</i> in year <i>t</i> (10 years following publication year) divided by the average citation rate of the papers published in the same year in the same discipline
$\ln(PublicfundingO)_{it}$	Natural logarithm of the three-year average up to year <i>t</i> of public sector funding for the purpose of operational costs and direct expenditures of research of researcher <i>i</i>
$\ln(PublicfundingI)_{it}$	Natural logarithm of the three-year average up to year <i>t</i> of public sector funding for the purpose of buying instruments for researcher <i>i</i>
$\ln(PrivatefundingO)_{it}$	Natural logarithm of the three-year average up to year <i>t</i> of private sector funding for the purpose of operational costs and direct expenditures of research of researcher <i>i</i>
$\ln(NFPfundingO)_{it}$	Natural logarithm of three-year average up to year <i>t</i> of funding from not-for-profit institutions (NFP) for the purpose of operational costs and direct expenditures of research of researcher <i>i</i>
$\ln(nbArticle)_{it}$	Natural logarithm of number of articles published in year <i>t</i> by researcher <i>i</i>
$\ln(nbAuthor)_{it}$	Natural logarithm of the three-year average up to year <i>t</i> of number of authors in the papers of researcher <i>i</i>
$\ln(Impactfactor)_{it}$	Natural logarithm of the five-year average up to year <i>t</i> of journal impact factor in which the scientist publishes
<i>dFemale_i</i>	Dummy variable taking the value 1 if the scientist is a woman and 0 otherwise
<i>Age_{it}</i>	Age of a researcher <i>i</i> at year <i>t</i>
<i>dUniv</i>	Dummy variables indicating the universities

Table 5 - Summary statistics (Number of observation = 39,911) – the variables are not summarized in logarithmic scale and they are raw amount²

<i>Variable</i>	<i>mean</i>	<i>Standard Deviation</i>	<i>min</i>	<i>max</i>
<i>nbCitation</i>	1.161812	1.820302	0	74.575
<i>dFemale</i>	0.231646	0.421889	0	1
<i>Tier1</i>	0.031571	0.174857	0	1
<i>Tier2</i>	0.024004	0.153063	0	1
<i>nbArticle</i>	3.257379	3.518717	1	85
<i>nbAuthor</i>	6.918378	49.73654	1	3174.5
<i>PublicfundingO</i>	105301.2	246244.1	0	1.59E+07
<i>PrivatefundingO</i>	19415.13	106610	0	6934758
<i>NFPfundingO</i>	19606.49	138335.3	0	1.12E+07
<i>Age</i>	50.12962	9.424855	14	92

² In some disciplines of Physics, there are many scientists involved in one project and therefore, the maximum for the number of authors is high.

Table 6 - Correlation table (all of them are significant at 1% level)

	$\ln(nbCitation)_i$	$Citavg3_i$	$dFemale$	$\ln(nbArticle)_i$	$\ln(PublicfundingO)_i$	$\ln(PrivatefundingO)_i$	$\ln(NFPfundingO)_i$	$\ln(Impactfactor)_i$
$\ln(nbCitation)_i$	1							
$Citavg3_i$	0.3453	1						
$dFemale$	-0.0311	-0.0589	1					
$\ln(nbArticle)_i$	0.2437	0.2787	-0.1015	1				
$\ln(PublicfundingO)_i$	0.0723	0.1084	-0.0312	0.237	1			
$\ln(PrivatefundingO)_i$	0.0789	0.1173	-0.1175	0.2299	0.1224	1		
$\ln(NFPfundingO)_i$	0.1053	0.129	-0.0203	0.2426	0.1773	0.2286	1	
$\ln(Impactfactor)_i$	0.4696	0.2516	-0.0314	0.2105	0.101	0.0468	0.1129	1
$\ln(nbAuthor)_i$	0.2856	0.2384	-0.0022	0.3091	0.0126	0.1624	0.2125	0.2366



Public-private collaboration and scientific impact: an analysis at the level of the individual researcher

Carter Bloch*, Thomas K. Ryan* and Jens Peter Andersen*

*carter.bloch@ps.au.dk; tryan@ps.au.dk; jpa@ps.au.dk

Danish Centre for Studies in Research and Research Policy, Aarhus University,
Bartholins Allé 7, 8000 Aarhus C., Denmark.

INTRODUCTION

This paper examines whether citation impact for individual researchers differs when collaborating with industry compared to work only involving academic researchers. To do this, we have identified a group of corresponding authors with addresses in Denmark with articles involving public-private collaboration for 2008-2010 and thereafter constructed a list of all articles authored by these researchers during the period 2006-2012.

One of the characteristics of the ‘entrepreneurial university’, and with it the promotion of public-private collaboration, are a greater utilization of academic research in business innovation (D’Este and Perkmann 2011), knowledge transfer and exchange, mutual learning, and a greater alignment of academic and business research. However, a concern here is that a shift in focus towards bridging academic and entrepreneurial research will detract from focus towards the type of ‘blue-skies’ research that often lies behind significant scientific breakthroughs (Etzkowitz 2003, Larsen 2011). In broader terms, this raises the question of how scientific impact is related to public-private co-authorships. Lebeau et al. (2008) examine the question using data for Canada and find that public-private collaborations receive on average higher (field normalized) citation counts than for university-only or industry-only papers.

However, in comparing university-only research with public-private collaboration, it is unclear whether we are comparing the same types of articles, for example concerning the nature of the research and the average performance of researchers involved. For example, the research of academics that engage in public-private collaboration may be more highly (or lowly) cited than those that do not. This paper is particularly focused on the role of individual performance by looking at a fixed group of researchers that have both engaged in public-private collaborations and university-only publications.

The main questions that we will examine here are the following:

- How does overall citation impact compare for academic papers and public-private collaborations when examining the same group of researchers?
- At the level of the individual researcher, is there a difference in citation impact for papers with and without public-private collaboration?

This paper is part of an ongoing project, Collaboration in Research, supported by the Danish Agency for Science, Technology and Innovation. The project utilizes publication and citation data to examine public-private collaboration and knowledge flows from university research to

the business sector in Denmark, and the impact of public-private collaboration on economic performance for Danish companies.

DATA AND METHOD

The sample examined in this paper consists of all corresponding authors with a Danish address with an article in 2008-2010 involving public-private collaboration, which amounts to a total of 798 researchers¹. For each of the 798 researchers, publication portfolios over the period 2006-2012 were collected in the Science and Social Science Citation Indices of Web of Science (WoS) using a name disambiguation algorithm developed by CWTS that has generally shown very high recall rates (90-95%) especially with sets of non-Asian author names and affiliations (Caron & van Eck, 2014)². Though, it is still important to note that this process is not perfect where both incomplete lists and false positives are possible. Of these 798 researchers, 46 had only one publication during the period, which precludes the possibility of any comparison across types of collaboration. Hence, these were removed from the sample.

In all, a total of 18,215 articles were identified over the period 2006-2012 for these 752 researchers. Table 1 shows the distribution of individuals according to number of publications.

Table 1. Distribution of individuals according to number publications

	2 to 5 pub.	6 to 10 pub.	11 to 19 pub.	20 or more pub.	Total	Number publications
Total	177	132	160	283	752	18,215

In general citation impact is typically much higher for papers involving international collaboration. Hence, we want to take account of international collaborations in the analysis here, and will utilize a classification of collaborations that divides articles into four mutually exclusive groups:

- Danish public research organization (university or government research institute), no industry co-author;
- Danish and international public research organizations, no industry co-author;
- Collaborations between Danish public research organizations and Danish companies (no international partners);
- Multiple collaborations - Danish public research organizations, at least one industry and at least one international partner

In this paper we first conduct an aggregated analysis, where we calculate the mean normalized citation score (MNCS), using the same approach as for the Leiden ranking (Waltman et al. 2012), for the entire subset of papers for each of these four types of collaboration. This subset of papers has in common that they include the same group of researchers as co-authors. However, it is clear that our approach here does not ensure that all co-authors are the same

¹ Due to data availability, it is more feasible to match other publications with the corresponding authors than with other authors.

² The collection of individual publication portfolios was performed by CWTS, Leiden University.

within each collaboration type, so this attempt to ‘level the field’ in terms of the researchers involved is only partially successful.

Secondly, we compare citation impact for the individual researcher. For each researcher covered in our sample, we calculate MNCS for the researcher’s publications in each of the four categories. We make pairwise comparisons of different types of collaboration in order to ensure that we are comparing the same group of researchers. So, when we for example compare MNCS for Danish public research papers with Danish public-private collaborations, we only include researchers that have publications in both of these groups. In order to be as inclusive as possible, we only require that one publication is needed in a category in order to calculate the MNCS for the individual researcher.

Finally, it is important to keep in mind that this data is highly skewed, which calls into question the validity of tests that assume that the data is normally distributed. While there are differing opinions on how extreme the skewness should be before precluding the use of t-tests in practice, it may be more appropriate to use non-parametric tests that essentially test whether overall distributions for two groups are equal. In the following, we report results of both tests.

RESULTS

The table below contains the mean and percentiles of impact for each collaboration type of all articles authored by the 752 authors in the dataset. The last column contains the average mean normalized citation score per researcher within each collaboration type.

Table 2. Mean Normalized Citation Score (MNCS) for different types of collaboration. Mean, median and percentile values for sample, and average values per individual. Period: 2006-2012.

	Sample mean	Median	75th pctl.	90th pctl.	99th pctl.	Avg. per indiv.	Mean for full sample 2006-2012
DK public research only	1.44	0.68	1.66	3.26	11.69	1.23	1.20
DK public research & DK industry	1.41	0.61	1.58	3.23	11.16	1.24	1.23
DK & Intl. public research	1.67	0.83	1.92	3.94	14.00	1.41	1.61
Intl. collab. incl. DK public research and industry partner	2.74	0.93	2.30	5.45	34.49	2.14	2.70
DK industry only	1.42	0.65	1.32	3.78	14.68	1.15	1.41
Total	1.67	0.75	1.80	3.66	15.06	1.46	1.50

The sample mean is the aggregate mean, calculated for all articles. For averages per individual, mean values are first calculated for each individual; thereafter the mean of individual averages is calculated.

As can be seen from table 2, median values are much lower than mean values, less than half of the mean in all cases. This indicates the large influence that the highest impact papers have

on overall averages. In the comparison of Danish public research papers and Danish public-private collaboration, the distributions are very similar, both for mean values, the median and also higher percentiles. Median values for papers with international collaboration are higher than for national collaborations, but in particular the median for international multi-partner collaborations is only slightly higher than for international public collaborations. The difference in median values between these two groups is only 0.10, while the difference in means is much larger at 1.07. Hence, the extremely high impact of a small group of papers appears to lie behind much of the difference in impact for international multi-partner collaborations compared to other papers.

Note again that the results above in table 2 are aggregated results for each type of collaboration. When we first compute MNCS for each individual researcher across these categories and thereafter calculate mean values per researcher, the resulting values of MNCS decline. Hence, average MNCS per researcher in this sample is 1.46 compared to an MNCS of 1.67 for all articles in the sample.

PAIRWISE COMPARISONS

The main results of this analysis are shown in table 3. The table shows the results of pairwise comparisons of MNCS within the four types of collaborations. Industry only papers are not included here as only a small number of researchers have these papers and a comparison would thus not be generalizable in any meaningful way.

Table 3. Comparing mean and median values per individual. Results of non-parametric and parametric tests.

	obs.	Mean	Median	Mann-Whitney (p-value)	T-test (p-value)
DK public research & DK industry	552	1.262	0.768		
DK public research only	552	1.198	0.964	0.003	0.231
DK public research only	504	1.273	1.015		
DK & Intl. public research	504	1.442	1.153	0.015	0.027
DK public research only	426	1.303	1.022		
Intl. collab. incl. DK public research and industry partner	426	2.248	1.164	0.021	0.002
DK public research & DK industry	475	1.285	0.772		
DK & Intl. public research	475	1.377	1.117	0.000	0.215
DK public research & DK industry	406	1.263	0.768		
Intl. collab. incl. DK public research and industry partner	406	2.283	1.15	0.000	0.002
DK & Intl. public research	416	1.449	1.190		
Intl. collab. incl. DK public research and industry partner	416	2.270	1.172	0.382	0.006

As noted above, comparisons are only made for researchers that have publications in both types of collaboration. For example, the first comparison is Danish public research papers vs. Danish public-private collaborations. Of the in all 752 researchers in the sample, 552 have at least one paper in each of these two groups. The table shows the average MNCS per researcher and median value of MNCS for this group of 552 researchers. We have conducted both standard t-tests and non-parametric Mann-Whitney tests on the data. P-values for the test statistics are shown in the table. The Mann-Whitney test is two-sided, while the t-test is one-sided (of whether the mean for the second group is significantly larger than the first).

Consider first the comparison of Danish public research papers (DK public only) and Danish public-private collaboration (DK PPC). Interestingly, the mean value of MNCS for DK PPC is larger than for DK public only, but the median value is actually smaller. Mean values are fairly close to one another, and the test of the equality of the means is not rejected. In contrast, the Mann-Whitney test is significant, implying that the distribution of MNCS values for DK PPC is lower than for DK public only.

The comparison for other types of collaboration is more straightforward. For example, in comparison of Danish public research with International research collaboration, international collaboration has both a significantly higher mean and the overall distribution is also significantly higher in terms of MNCS. The same result holds when comparing Danish public research with International collaborations with multiple partners.

In comparison of Danish public-private collaboration with international multi-partner collaboration, the results show clearly that citation impact for international public-private collaboration is higher than national public-private collaboration.

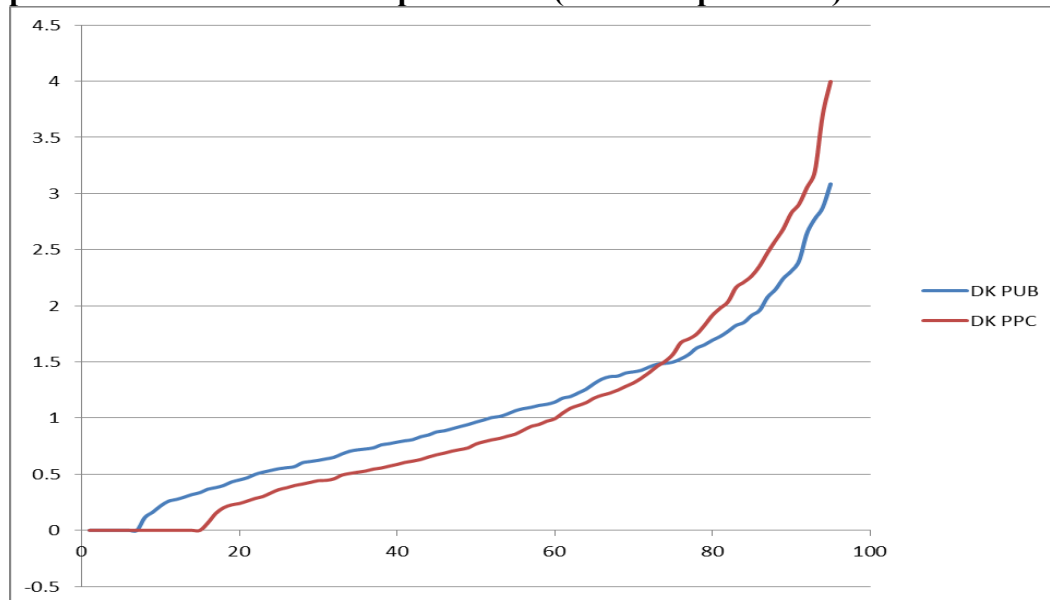
The final comparison in the table is between international collaboration only involving public research and multipartner collaborations that also involve public-private collaboration. Here there is a large difference in mean values, with average MNCS over 0.8 points higher for international collaborations involving industry. However, at the same time, median values for MNCS are almost the same. The statistical tests show that mean value of MNCS is significantly higher for multipartner, public-private collaborations, but that the overall distributions are not different from each other.

These results indicate that there may be some differences in the distributions of publications in relation to citation impact. To examine this further, we have compared distributions for national public research with national public-private collaborations, and international public research with international public-private collaborations. These distributions are shown in figures 1 and 2.

For Danish papers, citation impact (MNCS) is consistently higher for the bottom 75% of individuals for public research papers (where in particular, there is a much larger share with zero citations for PPC), while it is consistently lower for the top 25%. And if we consider only the top half of performers in terms of MNCS, the mean value is significantly higher for PPC. We can also calculate the difference for each individual, ie. the difference between MNCS for DK PPC and DK public research. When doing this, we find that MNCS for DK PPC is lower for 310 researchers and higher for 242 researchers. Finally, the mean value of the difference is 0.064 (not significantly different from zero). Overall MNCS is positively correlated with this difference (correlation coefficient), but not strongly so: 0.197 (ie a weak

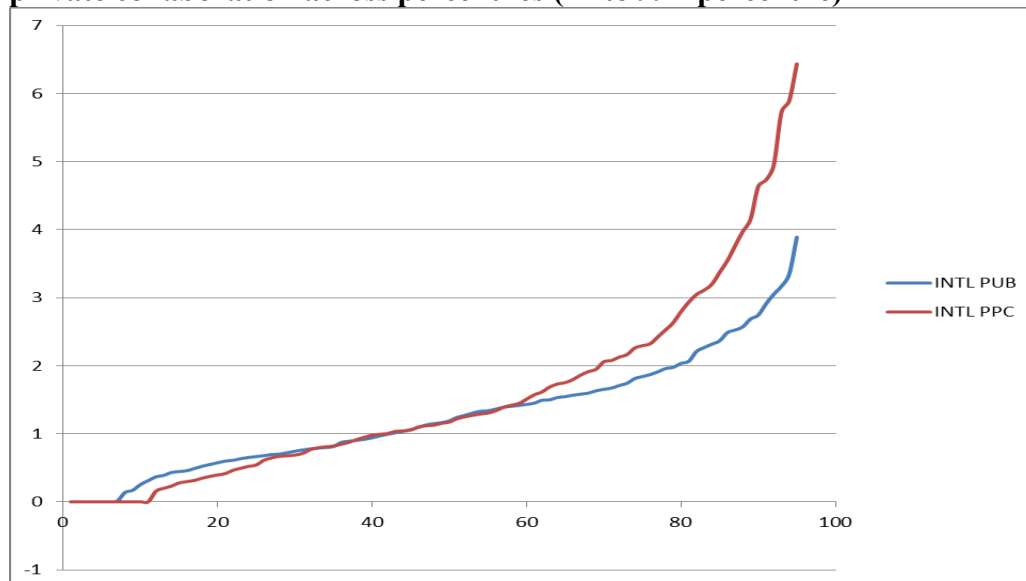
tendency that the higher the overall MNCS for the individual researcher, the higher the impact of their public-private collaborations is compared to university-only research)

Figure 3.4 - Distribution of MNCS for DK public research publications and DK public-private collaboration across percentiles (1st to 99th percentile)



For international papers, citation impact (MNCS) is higher for the bottom 25% of individuals for public research papers, equal in the middle, while PPC is much higher for the top 40%. If we consider only the top half of performers in terms of MNCS, the mean value is significantly higher for PPC.

Figure 3.5 - Distribution of MNCS for Intl public research publications and Intl public-private collaboration across percentiles (1st to 99th percentile)



As above, we can also calculate the difference for each individual, i.e. the difference between MNCS for international PPC and international public research. MNCS for international PPC is lower for 203 researchers and higher for 213 researchers. Mean value of the difference is

0.821 (significantly different from zero). Overall MNCS is positively correlated with this difference (correlation coefficient), more strongly than for DK papers: 0.523 (i.e. a strong tendency that the higher the overall MNCS for the researcher, the higher the impact of PPC is compared to university-only research)

CONCLUSIONS

In this paper we have compared the average impact (MNCS) of researchers that engage in both public research and public-private collaborations. We compared pairwise the average MNCS pr. researcher within four types of collaboration.

Results are slightly mixed concerning the comparison of citation impact for public research vs. public-private collaboration. This is particularly the case for papers only involving Danish authors. Average MNCS is higher for public-private collaboration, but this reflects large differences for a small share of researchers. For 56% of researchers, MNCS is higher for university-only publications. For international publications, average MNCS is much higher for public-private collaboration. Though, also here results are more mixed when looking at the full distribution, where MNCS for public-private collaboration is higher for only 51% of researchers. In general, it appears that differences between public-private collaboration and university-only publications are larger among those researchers with the highest citation impact overall. This is particularly the case for international papers. It is not clear from the analysis what lies behind this result. One possible explanation could be that highly cited researchers are better able to take advantage of the potential in public-private collaborations, or that they are more likely to enter into the more promising collaborative projects with industry. However, we are unable to examine whether this is the case based on the data here.

Our main focus here has been on individual averages - does the individual researcher typically have lower or higher citation impact when they collaborate with industry? However, if instead the focus is on the individual publication, then the aggregate figures we started with are more appropriate. There is an incredible degree of variation in the amount of publishing activity among researchers, with some producing a small handful of papers over the period and others producing over 100, or even over 300 in some cases. An analysis based on individual averages doesn't take this into account. We can see from table 3 above that, overall, there is no difference whatsoever in the mean value or distribution of MNCS when comparing national papers with and without public-private collaboration. For international papers in contrast, public-private collaborations have higher MNCS throughout, both in terms of the mean value and the overall distribution.

REFERENCES

- Caron, E. and van Eck, N. J. (2014) Large scale author name disambiguation using rule-based scoring and clustering. In Proceedings of the 19th International Conference of Science and Technology Indicators (pp. 79-86).
- D'Este, P. and M. Perkmann (2011). "Why do academics engage with industry? The entrepreneurial university and individual motivations." *Journal of Technology Transfer* 36(3): 316-339.
- Etzkowitz, H. (2003) Research groups as 'quasi-firms': The invention of the entrepreneurial university. *Research Policy*, 32(1), 109–121.
- Larsen, M.T. (2011) The implications of academic enterprise for public science: An overview of the empirical evidence. *Research Policy* 40: 6-19.
- Lebeau, L-M., Laframboise, M-C., Larivière, V. and Gingras, Y. (2008). The effect of university-industry collaboration on the scientific impact of publications: the Canadian case (1980-2005). *Research Evaluation*, 17(3): 227-232.
- Waltman, L., Calero-Medina, C., Noyons, E.C.M, Tijssen, R.J.W., van Eck, N.J., van Leeuwen, T.N., van Raan, A.F.J., Visser, M, and Wouters P. (2012) The Leiden ranking 2011/2012: Data collection, indicators, and interpretation. *Journal of the American Society for Information Science and Technology*, 63, 12, 2419-2432.



Survey on the Labour Market Position of PhD Graduates: Competence comparison and relation between PhD and current employment¹

Julia Heuritsch^{*}, Cathelijnn J. F. Waaijer^{**} and Inge C. M. van der Weijden^{***}

^{*} j.heuritsch@cwts.leidenuniv.nl

^{**} c.j.f.waaijer@cwts.leidenuniv.nl

^{***} i.c.m.van.der.weijden@cwts.leidenuniv.nl

Centre for Science and Technology Studies (CWTS). Faculty of Social and Behavioural Sciences, Leiden University, P.O. Box 905, Leiden, 2300 AX (The Netherlands)

ABSTRACT

We compared the skills PhD graduates acquired during their PhDs to the ones they need in their current job. We also studied the relation between PhD topic and content of the current job of recent PhD graduates. Data was collected in a survey of 1,133 respondents with a PhD from five Dutch universities between early 2008 and mid-2012. We show that scientific skills and independence are developed sufficiently during the PhD education, whereas PhDs are lacking in management and communication skills. These competence discrepancies were compared to the educational level required for the PhD holder's current job and the relatedness of the current job to the PhD topic.

INTRODUCTION

Traditionally, it is assumed that doctoral candidates devote most of their training and education to research, and that they represent the next generation of academic staff at universities. This point of view, however, does not account for the growing relevance of doctoral training towards a career in different types of labor markets (Kyvik & Olsen 2012, Hauss et al. 2015). Currently, career opportunities within academia have become uncertain and highly competitive, contracts being mostly temporary (Waaijer et al. 2015). In a survey of recent (1-6 years) PhDs from five Dutch universities, Maas et al. (2014) found that only 20% of the employed PhDs are in academic R&D, 60% in non-academic R&D, and the remaining 20% in non- R&D jobs, mostly in the private sector. Empirical evidence from Germany shows that one year after finishing the PhD about one quarter of doctorate holders in Germany are employed in academia (Kowalska 2013 in Hauss et al. 2015). Thus, though for many academia is the preferred sector of employment (de Goede et al. 2014, Sauermann & Roach 2012), an increasing number end up in other sectors. Particularly industry has become more attractive to young researchers. The reasons behind this seem to be strongly related to better career prospects and higher salaries in labor markets outside academia (e.g. Ostriker et al. 2011, Sauermann et al. 2012).

The large investment in time and money needed to train PhDs, makes it important that the skills obtained are those required in their future jobs, particularly now that these are

¹ Acknowledgements

We would like to thank Cornelis A. van Bochove (CWTS, The Netherlands) for his inspiring comments on this paper and Hans Sonneveld (Tilburg University, The Netherlands) for providing us with the survey sample on basis of which the data could be collected.

increasingly outside academia. In their training, as part of the process of doctoral socialization, PhD holders are equipped with scientific skills, personal effectiveness (in the remainder of the paper called “Independence”) and technological knowledge (Boosten et al. 2014). Do they need other skills in the non-academic sector, such as management, team and communication skills (dubbed “Management & Social skills”) more important? The aim of this paper is to study the relatedness of PhD holders’ current job with the topic of their PhD and the fit between the skills PhDs have acquired during their academic training and research, and the skills required in their job. Note that this does not necessarily mean an actual skill discrepancy, as PhDs could have developed some of the skills needed in their jobs before or after their PhD (or even during it, outside of their doctoral training). We will compare our results with those of Boosten et al. (2014) who conducted a similar survey, studying the careers of Flemish doctorate holders in 2010. Non-academic jobs often involve work below PhD level (Waaiker et al., in preparation), hence it is worthwhile to study whether doctorate holders are overqualified for their jobs and what the relation between the content of their current job and the topic of their PhD actually is. The results of this empirical study may guide politicians and practitioners in reforming doctoral education and doctoral programs.

METHODS

The survey was sent to the PhDs who obtained their degree between early 2008 and mid-2012 from five Dutch universities: three general universities (Utrecht, Leiden, and Rotterdam), one technical university (Delft) and one agricultural university (Wageningen). Data were obtained from 1,133 respondents (response just over 50%, Waaiker et al. 2015). Information was collected on the educational level required for respondents’ current job, on the extent to which the topic of their PhD is related to the content of their current job, to which extent skills were acquired during their PhD work and to which extent the same skills are required in their current job. Additional information was collected on many other factors, such as job and personal characteristics, perceptions and aspirations (Waaiker et al., in press).

The development and requirement of thirteen skills were measured in the survey which were chosen on basis of the relevance of the skill categories used by Boosten et al. (2014). Since there is a lot of covariance between these items, we condensed them by means of factor analysis, using the Principal Component Analysis (PCA) and Polychoric Correlations techniques, with FACTOR (Lorenzo-Seva & Ferrando 2006). This led to a structure matrix where the relation is given between the 13 initial components and 3 resultant factors: Scientific skills, Independence and Management & Social skills.

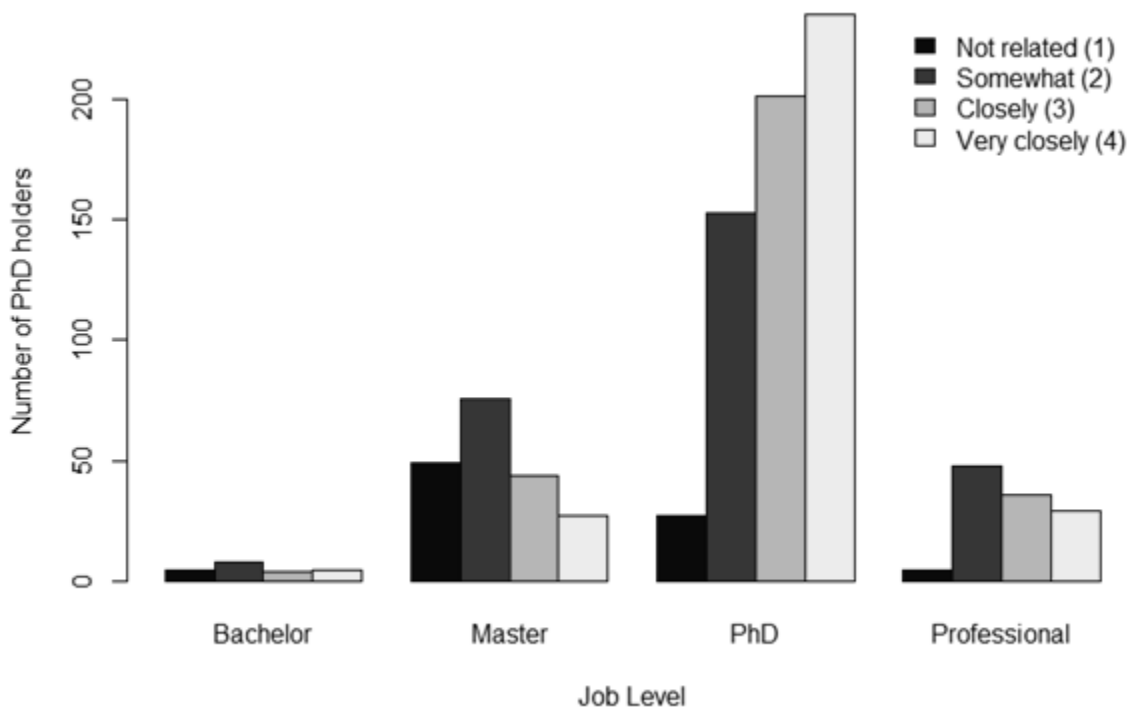
RESULTS

Relatedness job to PhD topic by Job Level

The results on job level and relatedness between the topic of the PhD and job content are presented in *Figure 1*. Note that “relatedness” is an ordinal variable but will be treated as a continuous one throughout the analysis for simplicity. Most PhD graduates (65%) work on PhD level and least on Bachelor level (2%). Of the PhD holders 21% work on Master level and a smaller fraction (12%) work on Professional level, which includes for example medical doctors. Therefore, nearly a quarter of PhDs work below PhD level, while the rest are employed at a level equivalent to or exceeding PhD level. As might be expected, PhD holders working on PhD level report that the content of their current job is highly related to their PhD

topic; the average relatedness equals 3.05 on a scale of 1 to 4. The content of the work of a PhD graduate working on professional level is closely related to their PhD topic as well, which is shown by an average relatedness of 2.75. Interestingly, the relatedness of PhD holders working on Bachelor level is higher (2.41) than those working on Master level (2.25), which might be because Master level jobs require more Management & Social skills than jobs on the other levels. This will be investigated further in *section 3.3*.

Figure 1: Job Level versus Relatedness

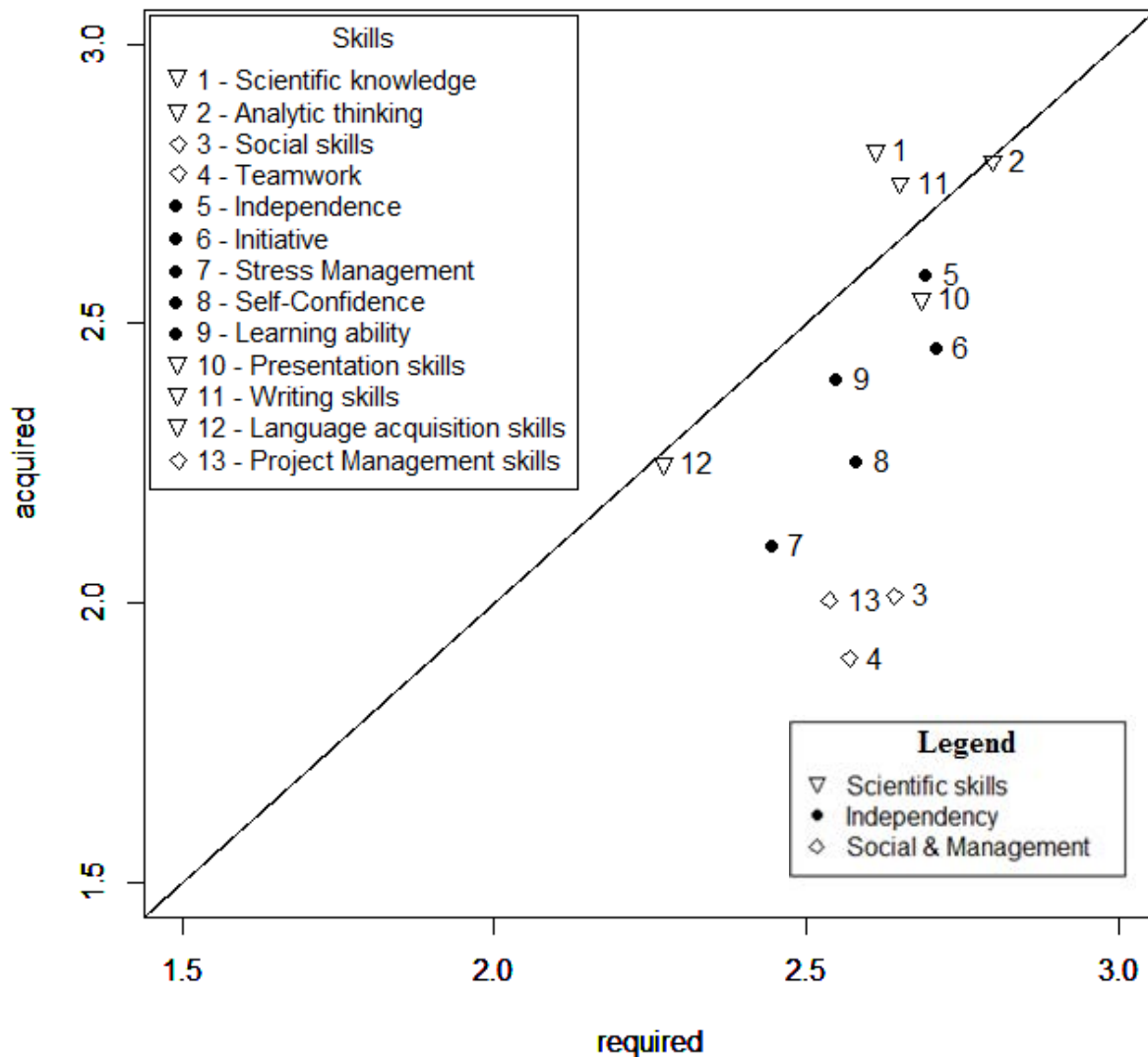


Skill discrepancy

Figure 2 plots the extent to which 13 skills were acquired during the PhD, against the extent to which those skills are required for the PhD holder's current job. The values correspond to the average of all respondents of the combined survey sample. They were asked to rate the extent to which skills were acquired and required on a scale of 1 to 3 (1: 'no', 2: 'yes, somewhat', 3: 'yes, very much'; an ordinal variable, treated as continuous for this analysis for simplicity). The three main sets of skills obtained by the factor analysis are Scientific skills (triangles), Independence (circles) and Management & Social skills (diamonds), which is demonstrated in both *Figure 2* and *Figure 3*.

As we would expect, Scientific skills are situated close to the diagonal, which means that they are sufficiently developed during the PhD. By contrast, PhD holders report that Management & Social skills were underdeveloped compared to what they would need for their current employment. Skills that represent the PhD holders' independence are located in between.

Figure 2: Skill discrepancy (13 categories)



Interestingly, the most underdeveloped Scientific skill is presenting (#10 in Figure 2), which makes sense as it involves social interaction and is therefore the scientific skill that is most closely related to the Management & Social skills.

The results compare very well to the conclusions drawn by Boosten et al. (2014): While employers highly appreciate the extensive Scientific skills and Independence of doctorate holders, for Management & Social skills, the gap is relatively wide. This may not be a problem if a required skill was acquired outside the PhD training. However, if that it is not the case and the skill is still underdeveloped, this poses a problem, because Management & Social skills are essential outside academia.

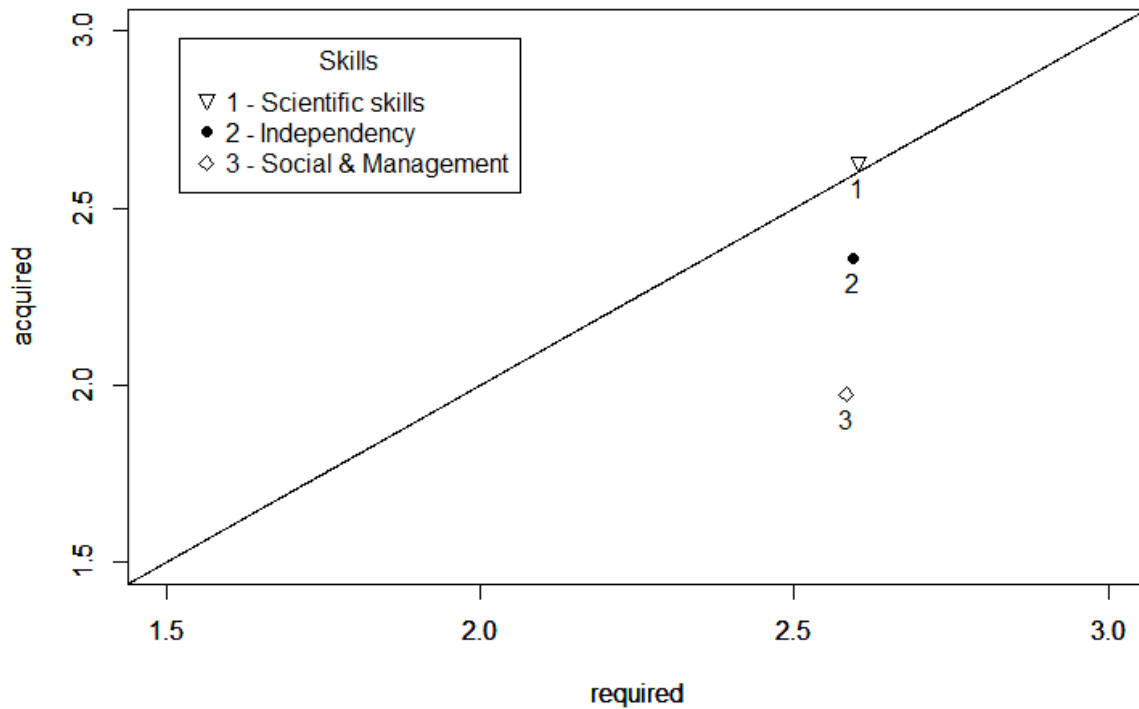
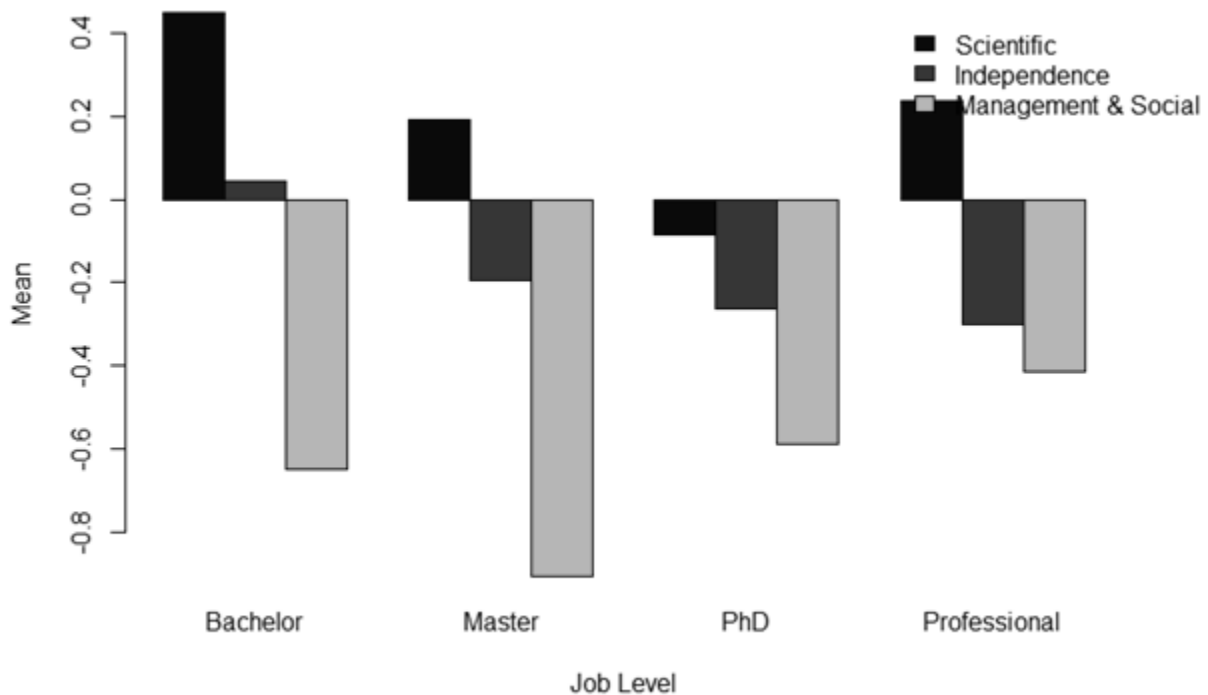
Figure 3: Skill discrepancy (3 categories)*Skill discrepancy by Job level*

Figure 4 compares the average values of skill discrepancy of the 3 skill factors, to the educational level that is required for the PhD graduates' current jobs. Positive discrepancy means that competences were acquired to a higher extent than required in the current job.

Interestingly, PhD holders working on PhD level show a relative lack in Scientific skills, while PhD graduates employed on all other levels perceive that those skills were developed more during their PhD study than required for the current job. This might be because most of those working on PhD level work in R&D (academic or non-academic), where their research work requires them to continue to develop their methodological knowledge and analytical skills. This is a subject for further study.

Figure 4: Skill discrepancy versus Job level

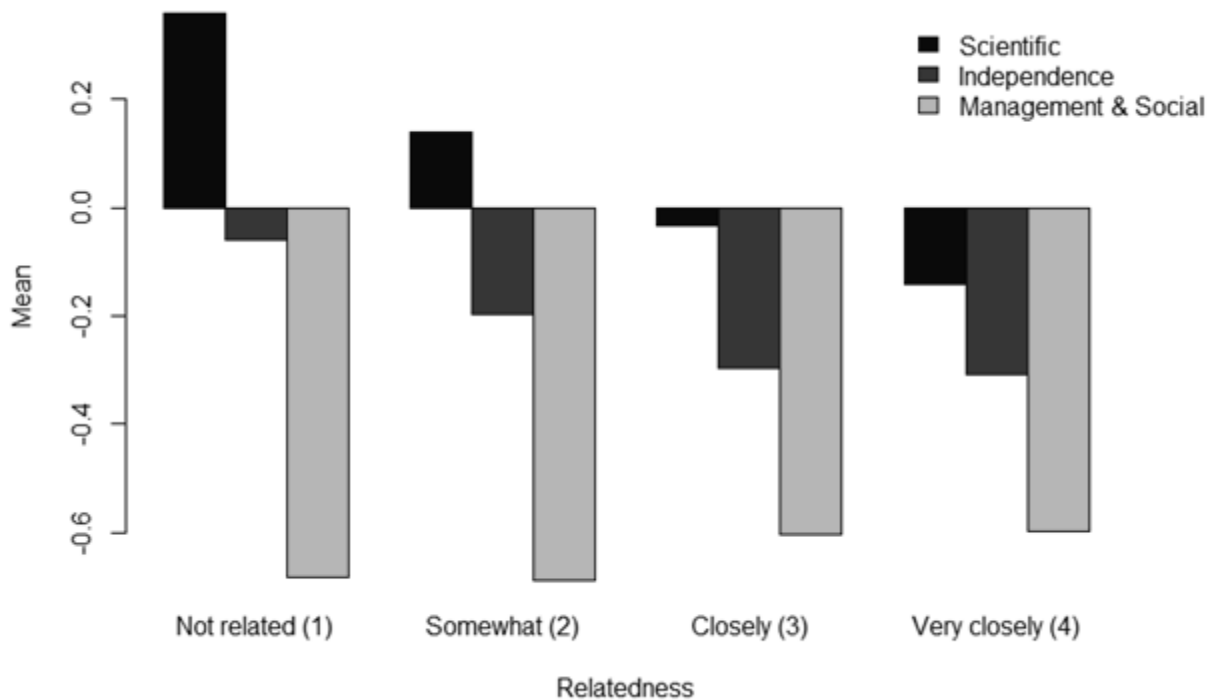
Management & Social skills are generally highly underdeveloped, which was already evident from *Figures 2 & 3*. Clearly, it is indeed work on the Master level that shows the highest discrepancy; this confirms the result in *section 3.1* on the relatedness of job content to PhD topic. As before, the reason for the higher requirement for that set of skills might be because jobs on Master level involve more management decisions than those at other employment levels. It would be worthwhile to investigate this further, for example by analyzing if those jobs include more supervisory tasks.

Skills associated with Independence are claimed to be overdeveloped only for PhD holders working on Bachelor level, which makes sense, as obtaining a Bachelor degree may not require the same level of personal effectiveness as the responsibility of conducting one's own research.

Skill discrepancy by Relatedness

Figure 5 compares the average values of skill discrepancy of the 3 skill factors, to the relatedness of the PhD topic to the current job content.

The closer related the field, the less adequate the acquired scientific skills are. This is in agreement with the observation that PhD holders working on PhD level, who reported the highest relatedness compared to the ones working on the other employment levels (*see Figure 1*), show the biggest lack in Scientific skills (*see Figure 4*).

Figure 5: Skill discrepancy versus Relatedness

The same decrease in the adequacy of acquired skills can be observed for the set of skills that reflect independence and personal effectiveness. On the one hand this is consistent with *Figure 4* that shows that PhD holders working on Bachelor level, reporting a low relatedness according to *Figure 1*, exhibit an excess in Independence. On the other hand, this finding is interesting because one of the goals of the PhD system is to train young researchers such that they can do research independently. Hence, the finding that the acquired independence skills are inadequate for the jobs most closely related to PhD work, raises the question whether that goal is achieved sufficiently. It would be interesting to investigate this phenomenon further.

The fact that the negative discrepancy of Management & Social skills decreases with relatedness reflects the finding reported above that those skills are needed most in Master level jobs (low relatedness according to *Figure 1*), and less in PhD level jobs (close relatedness). This is in agreement with the findings of Boosten et al. (2014).

CONCLUSION

Currently, there is a debate on whether the academic career system is efficient in attracting the most talented and fosters productivity. An essential part of this system is the PhD track, aimed to train young researchers to acquire a wide range of competences. We have shown that the development and use of those competences is related to the education required for the current job (Job level) and to the extent to which the topic of the PhD study is related to the current job of the PhD holder (relatedness).

In agreement with Boosten et al. (2014) we have shown that in skill training during the PhD, the emphasis is put on Scientific skills and Independence rather than Management & Social skills. This is reflected by the discrepancies between the skills acquired during the PhD work and those needed in current jobs. PhD graduates perceive their research skills to be almost

sufficiently developed in jobs closely related to PhD type work, but their personal effectiveness less so and their Management & Social skills even less. In the jobs not closely related to PhD type work, such as the non-research jobs that are becoming more prevalent among PhDs, the self-reported discrepancies between acquired and required competences are considerable: there is a redundancy of Scientific skills, a shortfall of Independence and a large deficit of Management & Social skills. Consequently, it appears that the current PhD system is inadequate from a labour market point of view. The fact that already, a lot of PhD holders are forced to work on Bachelor or Master level confirms this and points to an excess supply of PhDs.

Policy solutions could take either of two directions: the first would be to reduce the excess supply by decreasing the number of PhDs, e.g. by abolishing the PhD bonus universities now receive for every PhD graduate; clearly, the recently initiated experiments with a larger number of PhDs with a student instead of an employee status are a step in the opposite direction, as they are intended to *increase* the number of PhDs (Besluit experiment promotieonderwijs, 2016).

The second policy solution would be to alter the skill set acquired during PhD work, notably in the direction of more Management & Social skills, and of Independence. This might imply that the quantity of research required for a PhD would have to be decreased, and that more time would have to be spent on acquiring Management & Social skills as well as Independence skills, e.g. through substantial internships elsewhere. Similarly, it might be considered to shorten the PhD period, e.g. by reducing the current four-year, four peer-reviewed papers standard. The results of such policy measures could be visualized by repeating the survey regularly and deriving a time series of the skills discrepancy indicators we developed by our factor analysis.

This study will be continued by breaking down the results by sector of employment, as we can expect considerable differences in terms of competence requirements (Boosten et al. 2014). Thus the requirement for Management & Social skills is probably highest in business sectors. Do the PhD holders that work on Master level and report a substantial lack of those skills work in that kind of sectors or do they work in academia, performing supervisory tasks? Do most PhD holders who work on PhD level work in academic or perhaps in non-academic R&D?

Reflecting critically on the indicators developed in this study, it must be noted that our indicators are limited by the fact that they are based on perceptions of the PhDs themselves and on multiple choice questions that by their nature reduce reality to a stylized abstraction. Alternatively, one might try to obtain indicators on the skills of PhDs and the skills required for their post PhD jobs, by surveying their current employers. Furthermore, in future studies, both at employers and at PhD's, sets of open questions might be used to obtain a more richly textured picture of the experiences of PhDs.

REFERENCES

- Dutch Government. Besluit experiment promotieonderwijs (9th January 2016). Retrieved March 24, 2016 from: <http://wetten.overheid.nl/BWBR0037507/2016-01-09#Opschrift>
- Boosten, K., Vandeveld, K., Derycke, H., te Kaat, A. J. & Van Rossem, R. (2014). Careers of doctorate holders survey 2010. *Belgian Science Policy Office*, Brussels.
- De Goede M, Belder R, & De Jonge J (2014). Promoveren in Nederland: Motivatie en loopbaanverwachtingen van promovendi. *Rathenau Instituut*, The Hague.
- Hauss, K., Kaulisch, M. & Tesch, J. (2015). Against all odds: determinants of doctoral candidates' intention to enter academia in Germany. *International Journal for Researcher Development*, Vol. 6 Iss 2 pp. 122-143.
- Kowalska, K. (2013). B1.1 Trendaussagen auf Basis des Mikrozensus. Bundesbericht Wissenschaftlicher Nachwuchs 2013: Statistische Daten und Forschungsbefunde zu Promovierenden und Promovierten in Deutschland, 1st ed., W. Bertelsmann Verlag, Bielefeld, pp. 254-263.
- Kyvik, S. and Olsen, T.B. (2012). The relevance of doctoral training in different labour markets. *Journal of Education and Work*, Vol. 25 No. 2, pp. 205-224.
- Lorenzo-Seva, U., Ferrando, P.J. (2006). FACTOR: A computer program to fit the exploratory factor analysis model. *Behavior Research Methods*, Volume 38, Issue 1, pp 88-91
- Maas, B., Korvorst, M., van der Mooren, F. & Meijers, R. (2014). Careers of Doctorate Holders in the Netherlands. *Centraal Bureau voor de Statistiek*, The Netherlands
- Ostriker, J.P., Kuh, C.V. & Voytuk, J.A. (Eds.) (2011). A Data-Based Assessment of Research-Doctorate Programs in the United States: Committee on an Assessment of Research Doctorate Programs. *National Academies Press*, WA, DC.
- Sauermann, H., Roach, M. & Nunes Amaral, L.A. (2012). Science PhD career preferences: levels, changes, and advisor encouragement. *PLoS ONE*, Vol. 7 No. 5, pp. e36307.
- Sauermann, H., Roach, M. (2012). Science PhD Career Preferences: Levels, Changes, and Advisor Encouragement. *PLoS ONE*, 7(5), e36307.
- Waijjer CJF. (in press). Perceived career prospects and their influence on the sector of employment of recent PhD graduates. *Sci Publ Policy*.
- Waijjer CJF., Belder R., Sonneveld H., Van Bochove CA., & Van der Weijden ICM. (2015). Survey on the Labour Market Position of PhD Graduates: Development of a Novel Questionnaire. *CWTS Working Paper Series*, Centre for Science and Technology Studies, Leiden.



Beyond the indicators: formulation of the career strategies of scientists.¹

Éva PÁLINKÓ

palinko.eva@konyvtar.mta.hu

Department of Science Policy and Scientometrics, Library and Information Centre of the Hungarian Academy of Sciences, Arany János utca 1., Budapest, 1051, Hungary.

ABSTRACT

During the past years a comprehensive indicator-system has been developing in Hungary providing instant information for the purposes of evidence-based national science policy. The values in the indicator system are clear and comparable with international references. The suggested indices can point out important issues of the national R&D system and can contribute to raise relevant questions, however, it is emphasized that deeper analysis is needed during the decision making processes.

The aim of this article is to show the importance of such analyses based on the preliminary results of a career path research project among PhD holders.

The relevant science policy indicators show a favorable picture of the R&D system and the researchers' pay scale in Hungary by the wage premium of those with tertiary education compared with those with secondary education only. The deeper analysis shows a different picture. The first findings of the recent Career Path Monitor project among researchers at the Hungarian Academy of Sciences (HAS CPM) could provide additional concerns by understanding the background of the career decisions of the PhD holders deeper.

Data shows that some dimensions of satisfaction could determine the career-decisions of the examined scholars. The most problematic factor is wage owing to the characteristics of the public servant salary pay scale which is unfavourable for the young researchers at the beginning of their career. What is more it does not differentiate according to scientific performance, so the satisfaction with wage is usually low among young researchers. This problem makes researcher statuses less engaging causing drop in new supplies. Young researchers try to improve their financial circumstances in order to align their possibilities and their expectations based on their high qualification. SSH and STEM researchers have different optimization strategies, nevertheless both inhere severe risks (e.g. decline in scientific performance, brain drain).

INTRODUCTION

During the past years a comprehensive indicator-system has been developed in Hungary. The purpose of the system is to provide instant information for evidence-based national science policy. This suggested indicator system is being developed according to the strategic aims formulated in the current EU and national science policy documents regarding Hungary (Csité et al. 2013).

The values in the indicator system are clear and comparable with international references. The suggested indices can point out important issues of the national R&D system and can

¹ The study was funded under the IMPACT – EV FP7 collaborative grant agreement (n°613202) and the NKFI Hivatal_K_116163.

contribute to the emergence of relevant questions, however, it is emphasized that deeper analysis is needed during the decision making process.

The aim of this article is to show the importance of such analyses based on the preliminary results of a career path research project among PhD holders.

STRUCTURAL FRAMES: THE R&D SYSTEM IN HUNGARY

Research Policy Organisations in Hungary

In 1989, the democratic change in Hungary was slowly followed by the restructuring of different aspects of the polity and the society. As part of this process the former Soviet type of science organisations were replaced with a new structure based on the EU countries' standards and the elements of the previous national system.

In 1993, PhD was introduced replacing the former soviet type doctoral degree, the candidature. Universities regained the right to offer PhD programs and award PhD degree from the Scientific Qualification Committee, a semi-political body at the Hungarian Academy of Sciences during socialism.

Since the democratic transition the R&D structure has been changing continuously, but not only in connection with the structural change but because of the different governments. The changes of the organisations have been so intense that neither the researchers nor the organisations themselves could develop well and the system has been unpredictable for both of them.

The Hungarian R&D structure had effective institutions after the post-soviet transformation but these were not harmonised enough. Instead of having a focused and strong representation in policy making processes and its organisational structure, R&D function was dispersed in three ministries: Ministry for National Economy, Ministry of National Development, and Ministry of Human Capacities.

The National Research Development and Innovation Office have just been developed by the former president of the Hungarian Academy of Sciences at the Prime Minister's Office which may centralise the dispersed R&D deputy.

The governmental R&D background organisations at the ministries are focusing mostly on STEM research as the key of the economic growth, while paying less attention to SSH research.

Characteristics of the Research and Science Policy

Policy making processes and science and research policy documents were overly diversified, too partly due to the mentioned organisational status. There was a lack of unified terminology and well defined aims in the field of research and science policy making – because of the variety of different purposes and the variety of documents, aims were dispersed. The implementation of the written goals was very weak and inconsequent.

Since 2011 the importance of the reinforcement of the research community with young scientist have become better recognised (Report to the Hungarian National Assembly on the Activities of the HAS and on the General Situation of Hungarian Science, 2013).

The current science policy documents regarding Hungary show less, but clearer directions. The most recognizable aims are: to provide new supplies in R&D, strengthen the R&D infrastructure, improve scientific productivity, develop some internationally excellent research units, and to accelerate the collaboration between academic and business sector (National Research Development and Innovation Strategy 2013-2020; Csire et al 2013).

Research Institutions

The governmental, higher education and business enterprise sector are all active in the Hungarian R&D sector, but NGOs are not visible. The governmental research sphere almost entirely consists of the research organisations of the Hungarian Academy of Sciences. This is the most respectful platform of sciences in Hungary which has various research groups in all scientific fields. In higher education, researcher universities operate many research units in all scientific fields, too. The number of the higher educational research units has continuously been shrinking in the last decade. The number of the state-run research units, after a long stagnation in 2012, decreased drastically in connection with the reorganisation of the HAS institutions. By 2012 in Hungary the corporate sector operates the greatest number of research units (1,583) which has never happened before (KSH 2014).

Researchers

In 2014, 37,329 people were employed as researchers at different research organisations in Hungary. In 2012, the number of the FTE researchers per 1000 inhabitants in Hungary was 6,1 which is lower than the EU27 average 7,6 (EUROSTAT 2012).

After the democratic transformation the number of researcher positions drastically decreased, mostly in the industry. Around 1996, the correction began and has been continuous (KSH 2014). In 2006, the number of full time equivalent business enterprise researchers overtook both the number of academic and higher education researchers (KSH 2014). The trend of the last decade is that the traditionally relatively high number of academic, governmental and higher educational researchers is stagnating and the number of researchers in the business sector is growing. According to this the ratio, for the number of business enterprise researchers is at a very good level in a regional comparison, however, those who own a PhD degree are underrepresented in the business enterprise sector (EUROSTAT 2009). On the contrary their ratio in the government sector is very high in an international comparison – thanks to the traditionally strong academy in Hungary, the survivor of the former soviet-type science system.

Important Features of Academic Positions

The institution of tenure is common in Hungary, but the promotion had been incalculable for years after the transition, and still limited for young scholars.

In Hungary there is a linear relationship between seniority and pay in the public servant salary system for academic positions. Performance differences just have appeared sporadically between younger and older scholars.

Academic researchers are paid below the average compared to the researchers of the business sector in Hungary, and paid far below the average of the international (e.g. EU15) wage.

SCIENCE POLICY GOALS AND INDICATORS

In the actual EU and national science policy documents we can find obvious aims for Hungary. One of the most important science policy goals is to increase the number of highly qualified labour force (e.g. the number of researchers) (Partnership Agreement 2014-2020, National Research Development and Innovation Strategy 2013-2020, Csire et al. 2013). According to these and with the need for evidence – based science policy a comprehensive indicator-system has been developed to provide easily attainable information for the decision makers in R&D (Csire et al. 2013).

The developed indicators based on international standards are definite and comparable with international references what makes them very clear and useful. These indices can point out important issues of the national R&D system, and raise relevant questions, however, it is emphasized by the developer team, that in cases these values are not enough, deeper analysis is needed during the decision making processes.

The study of the indicator system identified 41 key national indicators as the most important ones in signalling the performance of the Hungarian R&D system in international comparison. The following 3 HR input indicators concern the mentioned growth aim conceived in the science policy documents:

Table 1: HR input indicators from the science policy indicator-system. Source: Csite et al. 2013. P. 43.

Indicator	Hungary	International reference	Source of reference
Number of researchers (FTE) per 1000 employed	6,1	7,6	EU27
Ratio of the people in tertiary education among young adults (age 18-24)	28,8%	30,4%	EU27
Ratio of PhD students among students in tertiary education	1,8%	2,3%	EU27

The values show that Hungary is lagging behind regarding the number of researchers per 1000 employed, and also regarding the ratio of people involved in higher education, however it is the ratio of PhD students compared to all students in tertiary education where the gap is the most significant. The indices reveal the narrow frames of R&D human resource supply which is opposite the current science policy goals.

Among the HR output indicators only two measures could have been identified in the mentioned study. Both measures show a favourable picture of the R&D system:

Table2: HR output indicators from the science policy indicator-system. Source: Csité et al. 2013. P. 43.

Indicator	Hungary	International reference	<i>Source of reference</i>
Wage premium of those with tertiary education, compared with those with secondary education only	72,0%	44%	<i>EU27</i>
Unemployment rate of those with tertiary education as a percentage of unemployment rate of those with secondary education only	42,1%	63,9%	<i>EU28</i>

By these values the wage premium of employees with a diploma in Hungary is much higher than the reference EU27 average. Interpreting only these indicators the researchers' pay scale seem to be adequately balanced and attractive enough for newcomers. At the level of the identified indicators only this could be recognised.

The first findings of our recent career path research project among scholars at the Hungarian Academy of Sciences could provide additional concerns to this science policy issue by understanding the background of the career decisions of the PhD holders deeper.

THEORETICAL BACKGROUND

The findings of this study are based on a longitudinal research project at the Hungarian Academy of Sciences (HAS) which follows the career path of scholars at the research institutions and evaluate the grants that are provided to them by .

The recent study focuses only on the *organisational* dimension of their career (Glaser, J., Laudel G. 2015) as well as the beliefs, and attitudes towards different organizational sectors and positions they could meet. As the project is an applied research project with a special focus on the organisational aspect of the career-path, it has severe theoretical limitations in case of attitudes.

In the literature of attitudes there is a great diversity of descriptions for the term, and methods for the measures. Definition of attitudes used in this concept is the one which is given by Fishbein and Ajzen (1975) as most investigators would probably agree with: "*a learned predisposition to respond in a consistently favourable or unfavourable manner with respect to a given object*" (Fishbein-Ajzen 1975: 10). They emphasize the distinction between *attitude* and other phenomena of the attitude area reserving *affect* as the most essential part of attitude. Other categories they point out are: "*cognition (beliefs, opinions), conations (behavioral intentions), and behavior (observed over acts)*" (Fishbein-Ajzen 1975:12).

During the different stages of the research project important information on the *beliefs, attitudes, intentions* and *behaviour* of scholars are gained, however, these results are not sufficient for a comprehensive, systematic attitude research. The results are relevant and implicate important concerns about the mentioned indicators.

METHODOLOGY

In the basic phase of the examination qualitative research was conducted among young researchers with PhD, widening the focus gradually, and finally detailed questionnaires were used. The research phases that have been carried out until now are the following:

- Career Path Research among Young PhD holders in Biology, a complex research with 11 semi-structured interviews, 2 narrative life story interviews, 2 focus groups and a small science field-specific survey (N=102);
- Career Path Research of Scholars with STEM (biology, chemistry physics, medicine) and SSH (sociology, political science, history, literature, linguistics) PhD in Hungary, with 30 semi-structured interviews, max. age 40 years.
- Career Path Monitor among research group leaders and members funded by the Lendület (Momentum) Programme of the Hungarian Academy of Sciences (assorted STEM and SSH fields, max. age 45 years); online survey (N=190).

Most findings of this article are based on the qualitative data from the early, explorative research phases. *Beliefs, attitudes and intentions* (Fishbein-Ajzen 1975.) towards the business sector had directly been examined in the very first phase with in-depth interviews and narrative life story interviews through the narratives of the interviewees. Later on the research, *organisational career-path* (Glaser, J., Laudel G. 2015) types were identified among scholars from different scientific fields based on their *actual* labour market *behaviour* (career sequences) unravelled from the semi structured interviews. Finally the researchers' *intentions* and *actual* labour market *behaviour* were tested in special labour market situations with questionnaires.

RESULTS

Main factors in the career path formulation

The analysis of the qualitative data identified three dimensions of the job satisfaction which can play a substantial role in forming the career paths of PhD holders. The dimensions of satisfaction proved to determine the career-decisions of the examined scholars are:

Tasks (Creativity, and meaningfulness)

Working environment (Motivation, inspiring colleagues and satisfying infrastructure)

Wage (Being able to live on without problems)

Highly qualified, motivated labour force try to keep these three dimensions at a consistently high level. The first dimension did not seem to be problematic in case of academics in Hungary: nearly all respondents like their tasks, feel that their job is meaningful and exciting. Regarding the second dimension there is a considerable variance of the answers: some researchers have reservations about the institutional circumstances at their institutions, others are satisfied. However, the third dimension proved to be severely problematic for many of the respondents.

Salaries are out of the focus of the international academic career research not being considered a measure of career success (Glaser, J., Laudel G. 2015). On the contrary in the case of Hungary wage proved to be important in career decisions and actual labour market behaviour. It is rooted in the characteristics of the Hungarian public servant pay scale, which is

unfavourable for young researchers at the beginning of their career and do not differentiate performance. Because of this structural circumstance the satisfaction with wage usually lags behind the level of the two above mentioned dimensions, causing inconsistency in the overall satisfaction with their academic statuses among young Hungarian scholars, which could result in severe frustration.

Findings show that young and postdoctoral researchers are eager to harmonise these dimensions, namely to improve their financial circumstances, in order to align their possibilities and their expectations which is based on their high qualification and motivational level. In different scientific fields they have different strategies for harmonising these factors, eliminating the inconsistency and getting over the frustration. Their *beliefs*, *attitudes*, *intentions* and their *actual behaviour* in the labour market show distinct strategies.

Career path strategies of researchers

SSH strategy

SSH careers are ‘*boundaryless*’ (Arthur and Rousseau 1996) in the meaning that SSH scholars are moving across the boundaries of different sectors, organisations and topics. The organisational sequences of the examined career-path stories draw out project-oriented ‘*multidirectional*’ careers (Baruch 2004) which are preferably based on a fix academic position. SSH researchers do not avoid business and enterprise sector. In their case the routine is to have complementary part-time jobs, consultative statuses, basic or applied research projects both in business enterprise and government sector besides their academic statuses.

Optimally, these projects connect to their own academic research topic. In this case these could improve their academic expertise and even their scientific performance directly. However, in many cases researchers have to work on many separate topics at the same time. It results in fragmented career span.

This strategy raises many questions. Could these researchers push a professional advantage or they simply miss some opportunities in their academic performance because of this strategy? A follow-up study should examine both the positive and negative effects on academic productivity of this fragmented career path structure and the impacts on the innovation of the business enterprise sector.

STEM strategy

STEM researchers usually don’t have complementary part time jobs or other “industrial” projects besides their academic positions, as it simply does not fit in their schedule. They have more ‘*linear*’ career paths (Baruch 2004). Their narratives show it is because they have to concentrate on their narrow field of research in order to keep up with their peers.

Nonetheless they react to the mentioned inconsistency, too. They have two main strategies: one is to apply for research grants in their field of interest which is a natural and useful part of their career-path, anyway. But the other one is dangerous, as it is to apply for a post-doctoral or even tenure status *abroad*.

The most important finding is that Hungarian STEM scholars prefer foreign academic positions to business and enterprises researcher jobs in Hungary. Both qualitative and quantitative results about their *intentions* and their labour market *behaviour* underpin that most of the STEM researchers would leave the country instead of changing sector inside Hungary.

In the background of this phenomenon we have found different factors. The negative *beliefs, and attitudes* towards the business enterprise sector were one: by the results of the qualitative data, the *attitudes, and beliefs* of young STEM researchers formed strong and commonly shared negative stereotype, saying that business enterprise jobs are ‘monotonous’, ‘dull’, and ‘boring’, ‘not requiring any creativity’. This proved to be very important as ‘creativity’ and ‘exciting work’ with ‘autonomy’ were the most important positive principles they attributed to their academic researcher jobs. These negative beliefs and attitudes towards the business sector jobs may root partly in the traditional intellectual role interpretation according to the common values in higher education (Palló 2009).

What is more, they fear that changing the academy to business sector means the end of their scientific career because of the limited publishing possibilities. They choose the opportunities which could keep them in their scientific career paths without breaking its span – this is exactly what they are optimising for.

Another important background factor emerging from the career narratives is that the reference group regarding wage for these internationally mobile young scientists is usually the international or the EU15 scholars’ community and its’ attainable standard of living. They compare their financial possibilities to the Western European counterparts and not to Hungarian with a secondary education only – and this is the point where the mentioned indicators prove not to be enough in themselves.

All these factors regarding the background of the career decisions of the examined scholars must be important as underlying causes of the high level of brain drain among STEM researchers in Hungary. According to a calculation (Csanády-Személyi 2006), one in every four fresh graduate with a diploma in science leaves the country. This rate for PhD holders is even higher (Csanády-Kmetty-Kucsera-Személyi-Tarján 2008).

The main question is, under which circumstances would they be willing to come back, or stay. Under what structural circumstances can they harmonise better the mentioned factors in order to gain satisfaction at their academic positions in Hungary? Could the business enterprise sector in Hungary offer any remedy for this brain-drain problem?

CONCLUSION

It has always been clear, that the currently used indicators don’t provide enough information regarding complex issues in the field of science policy. The deeper analysis has shown that policy makers have to acquire more information in certain cases, e.g. the needs of highly qualified and motivated young researchers desperate for academic research, and their concerns with the actual regulations of the public servant salary scale.

Further research should focus on the possible changes in the public servant salary system relevant for researchers (e.g. the premium by performance), and even the role of the business sector in keeping the most creative labour force in the country.

REFERENCES

- Arthur, M.B., Rousseau, D.M. (1996): *The boundaryless career: A new employment principle for a new organizational era*. New York: Oxford University Press.
- Baruch Y (2004): Transforming Career: From Linear to Multidirectional Career Paths. Organizational and Individual Perspectives. in: *Career Development International* 9 (1): 58-73.
- Csanády Márton – Személyi László (2006). Brain Drain. Közelkép a diplomás magyarokról. in: *Századvég*. 41, 3, 79–122.
- Csanády-Kmetty-Kucsera-Személyi-Tarján (2008). A magyar képzett migráció a rendszerváltás óta. in: *Magyar Tudomány* 2008/5.
- Csite, A., Major, K., Mike, K., Remete, Zs., Zsár, V., (2013) Stratégiai irányok meghatározása, javaslat a tudománypolitikai stratégiára. (Identifying strategic directions, suggestion for a science policy strategy) HÉTFA
- Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley
- Glaser, J., Laudel G. (2015). The Three Careers of an Academic. Discussion paper Nr. 35/2015 09/2015.
- Palló, G (2009): A fiatal kutatók tevékenységének erőtere: A magyar tudományrendszer vázlatja *Magyar tudomány* 2009 (2): 11 – 22
- National Research Development and Innovation Strategy (Befektetés a jövőbe) 2013-2020 Partnership Agreement 2014-2020
- Policy Documents and Statistics:
- Doctorate Holders Working as Researchers by Sector of Performance (% of Total Doctorate Holders Employed as Researchers) EUROSTAT (2009)
- http://ec.europa.eu/eurostat/statistics-explained/index.php/File:Doctorate_holders_working_as_researchers_by_sector_of_performance_%28%25_of_total_doctorate_holders_employed_as_researchers%29,_2009.PNG
- Report to the Hungarian National Assembly on the Activities of the HAS and on the General Situation of Hungarian Science (2013)
- <http://www.parlament.hu/irom39/13472/13472.pdf>
- “Investment into the Future” National Research-development and Innovation Strategy (2013-2020) “Befektetés a jövőbe” Nemzeti Kutatás-fejlesztési és Innovációs Stratégia
- http://2010-2014.kormany.hu/download/b/35/f0000/06_11_NGM%20KFI%20strat%C3%A9gia_Kozlonyhoz.pdf
- A kutató-fejlesztő helyek kutatóinak számított létszáma, KSH (2014)
- https://www.ksh.hu/docs/hun/xstadat/xstadat_eves/i_ohk002f.html



Exploring predictors of scientific performance with decision tree analysis: The case of research excellence in early career mathematics

Jonas Lindahl*

*jonas.lindahl@umu.se

Inforsk (Department of sociology), Umeå University, Umeå, SE-901 87 (Sweden)

ABSTRACT

The purpose of this study was (1) to introduce the exploratory method of decision tree analysis as a complementary alternative to current confirmatory methods used in scientometric prediction studies of research performance; and (2) as an illustrative case, to explore predictors of future research excellence at the individual level among 493 early career mathematicians in the sub-field of number theory between 1999 and 2010. A conceptual introduction to decision tree analysis is provided including an overview of the main steps of the tree-building algorithm and the statistical method of cross-validation used to evaluate the performance of decision tree models. A decision tree analysis of 493 mathematicians was conducted to find useful predictors and important relationships between variables in the context of predicting research excellence. The results suggest that the number of prestige journal publications and a topically diverse output are important predictors of future research excellence. Researchers with no prestige journal publications are very unlikely to produce excellent research. Limitations of decision tree analysis are discussed.

INTRODUCTION

Bibliometric indicators are increasingly used as decision support tools in academia (Abramo, Cicero, & D'Angelo, 2013). Indicators that are used as decision support tools should satisfy at least two basic assumptions: (1) indicators of past performance should be able to predict future scientific performance (e.g., Danell, 2011); and (2) indicators should be free from inherent biases (e.g., Moed, 2005). These assumptions have been tested in numerous scientometric prediction studies at different levels of aggregation (e.g., Jensen, Rouquier, & Croissant, 2009; Penner, Pan, Petersen, Kaski, & Fortunato, 2013; Dubois, Rochet, & Schlenker, 2014; Havemann & Larsen, 2015).

Most scientometric prediction studies are conducted in a confirmatory framework based on testing hypothesis. In this study I suggest the framework of exploratory data analysis (EDA), and specifically decision trees, as an underutilized source of methods that could complement the dominating confirmatory framework (Strobl, Malley, Tutz, & Maxwell, Scott, 2009; King, & Resick, 2014).

Decision trees has a number of desirable features in the context of predicting research performance from the micro to the macro level. Decision trees (King & Resick, 2014): (a) are non-parametric; (b) are flexible and can handle many different variable types; (c) can identify non-linear relationships; (d) can identify useful predictors; (e) show interactions between

predictors without the need to specify these in advance; and (f) are transparent, intuitive to interpret, and can be used as visual tools to inform decision making.

Publication praxis and citation behaviours differ between scientific fields (Moed, 2005). Decision trees provide an opportunity to explore research fields with few prior assumptions. Such data driven exploration can generate new hypothesis to test with confirmatory methods. Decision tree analysis may prove particularly useful for the study of peripheral and emerging areas of research where knowledge is scarce to begin with.

The purpose of this study is (1) to introduce the exploratory method of decision tree analysis as a complementary alternative to current methods used in scientometric prediction studies of research performance; and, as an illustrative case, (2) to identify important predictors, interactions between predictors, and the effect of combinations of publication track record characteristics to predict research excellence at the individual level among 493 early career mathematicians in the sub-field of number theory between 1999 and 2010.

METHOD

Data collection

The dataset consisted of article publication track records of 493 authors in number theory. The authors were selected on the basis of (1) at least one published article in class 11 (i.e., Number theory) in the Mathematics Subject Classification (MSC) scheme between 1999 and 2003; (2) an article publication career of \geq eight years; and (3) that the share of articles belonging to the MSC class in the track record of an author was \geq than the share of any other MSC class found in that authors track record.

Publications ($N=4654$) was retrieved from the MathSciNet (MSN) database and matched against publications indexed in the citation indices available through Web of Science (WoS) Core Collection to obtain citation data. The final dataset retrieved from WoS consisted of 2975 articles, reviews, notes and letters.

Design and variables

The design comprise two time periods: Period 1 (P1) and Period 2 (P2). P1 is the first four years in the publication career of an author. P2 is the fifth to the eighth year. The publication career of an author begin with the first MSN article publication.

The response variable consisted of a binary variable indicating if an author publish \geq one article in P2 that can be considered excellent (i.e., an excellent researcher). An article is defined as excellent if it has a document type, publication year, and field normalized citation score (FNCS) adjusted for multiple Web of Science Categories \geq the 90th percentile given a reference set (Lundberg, 2007). I used the article FNCSs of the publications of the 493 authors in P2 as a reference set to calculate the percentile.

Seven predictors were used in the analyses. Predictor:

- (1) address publication rate in p1 (coding: P) and consist of the number of MSN articles in P1;
- (2) address journal prestige and consist of the number of articles published in journals with a SNIP value \geq the 75th percentile in the CWTS Journal Indicators list (CWTS,

- 2015; Waltman, Van Eck, Van Leeuwen, & Visser, 2012) in the publication year of the article in P1 (coding: *Top_Jour*);
- (3) address collaboration and consist of the average number of authors per publication during P1 (coding: *Avg_Co_Au*);
 - (4) address topical diversity (coding: *Topic_Div*) and consist of the number of different MSC classes an author has published in during P1;
 - (5) address mobility in P1 and consist of the number of MSN publications at different universities (coding: *Mob_Univ*);
 - (6) address university prestige (coding: *Top_Univ*) and is a binary predictor indicating if an author has published ≥ 1 publication at a top university in P1;
 - (7) is a binary predictor and address whether an author has published $\geq 50\%$ of her/his output at institutions located in English speaking countries during P1 (coding: *English*).

Decision tree analysis

A decision tree is built by an algorithm that successively split the initial dataset into smaller sub-groups based on splitting rules (King, & Resick, 2014). When the predictors and the response variable are chosen the decision tree is built in three main steps.

In step one all predictors are evaluated to find the “best” binary split. The decision tree algorithm starts with the total dataset. The best binary split is a cut-off threshold among the values in the chosen predictor. When the cut-off is determined the dataset is divided in two sub-groups. The goal of splitting is to assign authors with similar values in the response variable in the same group so that the two sub-groups are more homogenous than the previous group (King, & Resick, 2014).

In step two the splitting procedure in step one is performed on the total dataset and two sub-groups are created. Each sub-group is treated as a new dataset; a cut-off for the best binary split is determined and two sub-groups is created. Successively the decision tree algorithm creates smaller and more homogenous groups. All predictors are evaluated for each potential split (King, & Resick, 2014).

In the third step the splitting procedure is ended by some stopping criteria. The definition of the best split and the stopping criteria depends on the decision tree algorithm. In this study I used an implementation of the algorithm for conditional inference trees (Hothorn, Hornik, & Zeileis, 2006) in the Party package available through R (R Core Team, 2015). With this algorithm the best binary split is determined by testing the global null hypothesis of independence between each of the predictors and the response variable with permutation tests and further between the each possible binary sub-set of the chosen predictor and the response (Hothorn, Hornik & Zeileis, 2006). The predictor and cut-off threshold resulting in the strongest association with the response variable (i.e., lowest p value) are chosen for the split. The splitting procedure stop when the global null hypothesis of independence between all possible combinations of the predictors and the response are rejected at some pre-specified level of alpha (e.g., 0.05; Hothorn, Hornik & Zeileis, 2006). Predictors that has not been chosen for a split when the splitting procedure stops are not included in the tree model.

The performance of decision tree models are usually evaluated with the statistical method of cross-validation (Maimon & Rokach, 2008). The basic principle of cross-validation is to split

This work is licensed under a Creative Commons License: Attribution-NonCommercial-NoDerivatives 4.0 International.

the initial dataset into a training set and a test set. The decision tree model is first trained on the training set and then validated on the test set. The purpose of cross-validation is to evaluate the generalizability of the decision tree model by fitting the trained model on new data (i.e., the test set) (Maimon & Rokach, 2008). By fitting the trained model on new data a more realistic performance measure can be obtained since models tend to have a better fit on the training set due to overfitting, compared to a new dataset that is sampled from the same population as the training set (Maimon & Rokach, 2008).

In this study I tested the model with the commonly used method of 10-fold cross-validation, where the dataset is split into 10 equally sized and non-overlapping folds (i.e., sub-groups) (Maimon & Rokach, 2008). The cross-validation consists of ten iterations. In each iteration one fold is used as the test set and the remaining nine folds are used as the training set. In each iteration some appropriate metric of model performance is calculated. As a result of the 10-fold cross-validation the ten values of the chosen metric are averaged to produce a single cross-validated performance measure (Maimon & Rokach, 2008).

RESULTS

Decision tree analysis: Predictors of research excellence among early career mathematicians

Figure 1 depicts a decision tree which consist of a single root node (oval) at the top, a number of internal nodes (ovals) and several terminal nodes (bar charts) at the bottom. The nodes in the tree is connected by branches. Each split is represented by a predictor label visible in the node denoting which predictor was used for the split (e.g., *Top_Jour*). The value at which the best split occurred in the predictor is placed along the branches between nodes. The splitting procedure stops at the terminal nodes. Each terminal node provide a bar chart indicating the proportion of authors in each class and the *n* of authors in that group.

32 authors had missing values on the affiliation based predictors and was excluded from the analysis. The analyses was performed with 461 authors. Of the 461 authors, 71 was defined as excellent (i.e., incidence=15.4%). The binary response variable and seven predictors was used as input for the decision tree depicted in Figure 1: *P*; *Top_Jour*; *Avg_Co_Au*; *Topic_Div*; *Mob_Univ*; *Top_Univ*; *English*. However, only three of these predictors, *Top_Jour*, *Topic_Div*, and *English* was actually used in the tree. This indicates that *P*, *Avg_Co_Au*, *Mob_Univ*, and *Top_Univ*, did not contribute to the model.

I used the *10-fold cross validated area under the receiver operating characteristic (ROC) curve (AUC)* to evaluate the prediction accuracy of the decision tree (King & Resick, 2014). The AUC is not sensitive to skewed class distribution which make it an appropriate metric in the context of bibliometric data (Maimon & Rokach, 2008). The decision tree model had a cross-validated AUC of 0.73, indicating acceptable discrimination between excellent authors and non-excellent authors according to the rule of thumb interpretation of AUC-values suggested by Hosmer and Lemeshow (2000).

Figure 1. Decision tree consisting of three predictors of research excellence among 461 mathematicians

Predicting research excellence as combinations of publication track record characteristics

Each author follow a path through the tree and end up in a terminal node (Figure 1). The path through the tree reveal the combinations of publication track record characteristics (as defined

This work is licensed under a Creative Commons License: Attribution-NonCommercial-NoDerivatives 4.0 International.

by the predictors and the predictor values) that is required for an author to end up in that particular terminal node.

At the root node, all authors (N=461) are evaluated for a potential split (node 1). Authors with ≤ 3 in *Top_Jour* follow the left branch and authors with > 3 in *Top_Jour* follow the right branch. The group of authors with > 3 in *Top_Jour* is further split on the predictor *Topic_Div*. Authors with a *Topic_Div* of ≤ 3 end up in terminal node 10, a group consisting of 8 authors of which 12.5% is excellent (Table 1). The group of authors with a topical diversity > 3 end up in terminal node 11, a group consisting of 32 authors of which 62.5% is excellent (Table 1). The appearance of *Topic_Div* (node 9) in the branch to the right of *Top_Jour* (node 1) but not to the left represents an interaction. Thus, *Topic_Div* has an effect on future research excellence at high levels of *Top_Jour*. *Top_Jour* and *Topic_Div* seem to be the best predictors of future research excellence.

Table 1. Showing combinations of predictors and predictor values required to end up in a particular terminal node.

TN*	Combinations of publication track record characteristics	% of excellent authors	n of authors	% of total authors
11	<i>Top_Jour</i> > 3 ; <i>Topic_Div</i> > 3	62.5%	32	6.9%
10	<i>Top_Jour</i> > 3 ; <i>Topic_Div</i> ≤ 3	12.5%	8	1.7%
6	<i>Top_Jour</i> ≤ 3 ; <i>Top_Jour</i> > 0 ; <i>Topic_Div</i> ≤ 3 ; <i>English</i> = Yes	26.8%	71	15.4%
7	<i>Top_Jour</i> ≤ 3 ; <i>Top_Jour</i> > 0 ; <i>Topic_Div</i> ≤ 3 ; <i>English</i> = No	6.8%	103	22.3%
8	<i>Top_Jour</i> ≤ 3 ; <i>Top_Jour</i> > 0 ; <i>Topic_Div</i> > 3	26.2%	61	13.2%
3	<i>Top_Jour</i> ≤ 3 ; <i>Top_Jour</i> ≤ 0	4.3%	186	40.3%

* Terminal Node

Authors following the left branch with ≤ 3 but > 0 in *Top_Jour* is further split on the predictor *Topic_Div*. Authors with > 3 in *Topic_Div* end up in terminal node 8, a group consisting of 61 authors of which 26.2% is excellent (Table 1). Authors with a *Topic_Div* value ≤ 3 is further split on the binary predictor *English*. An author with $< 50\%$ of the publication output at an institution in an English speaking country end up in the group represented by terminal node 7, of which 12.5% is excellent (Table 1). Authors with $\geq 50\%$ end up in terminal node 8 of which 26.7% is excellent.

The predictor *English* is important among authors with > 0 but ≤ 3 in *Top_Jour* and ≤ 3 in *Topic_Div* (Table 1). This interaction indicate that publishing in an English speaking environment early in the career increase the probability of producing excellent research in P2 at low levels of *Top_Jour* and *Topic_Div*.

The group of authors with ≤ 3 papers in top journals in the first split and ≤ 0 papers in top journals in the second split, end up in terminal node 3, a group consisting of 186 authors of which 4.3% is excellent. At this level of *Top_Jour*, the predictors *Topic_Div* and *English* has no effect on the outcome.

DISCUSSION AND CONCLUSION

This work is licensed under a Creative Commons License: Attribution-NonCommercial-NoDerivatives 4.0 International.

Decision trees can identify useful predictors and important relationships between variables without the need to specify a model a priori. In the case of 493 mathematicians in number theory seven predictors of research excellence was used as input, but only three, *Top_Jour*, *Topic_Div*, and *English*, was included in the tree. These results suggest that early career publication strategies where prestige journal publications and a topically diverse output is important for the production of future excellent research (as defined in this study). Previous research has shown that topical diversity has a positive effect on productivity (Dubois, Rochet, & Schlenker, 2014).

A particularly useful feature of decision trees is their ability to reveal interactions between predictors without specifying these in advance. An interesting interaction between the predictors, *Top_Jour*, *Topic_Div*, and *English*, was identified. At lower levels of *Top_Jour* and *Topic_Div* the predictor *English* had an effect on the response variable indicating a compensatory effect of early career publication activity in an English speaking environments. This interaction affected 37.7% of the included authors. Since the response variable used as a proxy for research excellence in this study is based on WoS data, the effect of *English* on the outcome may be a consequence of the well-known English language bias inherent in WoS (Moed, 2005).

Another useful feature of decision trees is the ability to show how different combinations of publication track record characteristics affect the outcome. Results show that authors with at least four publications in top journals and publications in at least four different subject areas have the highest probability to produce excellent research in P2. High topical diversity also has an effect on lower levels of *Top_Jour*. Authors with no publications in top journals are very unlikely to produce excellent research in P2. The information provided can be used to explore which combinations of publication characteristics that would be prioritized given some research policy selection criteria (e.g., a citation based indicator).

One limitation with decision trees is a tendency towards instability (i.e., small changes in the data can cause significant changes in the tree structure) in some situations (e.g., small sample size; overfitted trees) (King & Resick, 2014). Given this limitation decision trees are suitable as a complementary method to identify relationships between predictors that can be further tested with confirmatory methods.

REFERENCES

- Abramo, G, Cicero, T, & D'Angelo, A. (2013). Are the authors of highly cited articles also the most productive ones? *Journal of Informetrics*, 8(1), 89-97.
- Danell, R. (2011). Can the quality of scientific work be predicted using information on the author's track record? *Journal Of The American Society For Information Science And Technology*, 62(1), 50-60.
- Dubois, P., Rochet, J., & Schlenker, J. (2014). Productivity and mobility in academic research: Evidence from mathematicians. *Scientometrics*, 98(3), 1669-1701.
- Havemann, F., & Larsen, B. (2015). Bibliometric indicators of young authors in astrophysics: Can later stars be predicted? *Scientometrics*, 102(2), 1413-1434.

Hosmer, D., & Lemeshow, S. (2000). *Applied logistic regression*. New York: Wiley.

Jensen, P., Rouquier, J. B., & Croissant, Y. (2009). Testing bibliometric indicators by their prediction of scientists promotions. *Scientometrics*, 78(3), 467-479.

King, M., & Resick, P. (2014). Data Mining in Psychological Treatment Research: A Primer on Classification and Regression Trees. *Journal of Consulting and Clinical Psychology*, 82(5), 895-905.

Lundberg, J. (2007). Lifting the crown—citation z-score. *Journal of Informetrics*, 1(2), 145-154.

Maimon, O., & Rokach, L. (2008). *Data Mining With Decision Trees: Theory and Applications*. London: World Scientific.

Moed, H. (2005). *Citation analysis in research evaluation*. Dordrecht: Springer.

Penner, O., Pan, R. K., Petersen, A. M., Kaski, K., & Fortunato, S. (2013). On the Predictability of Future Impact in Science. *Scientific Reports*, 3, 3052.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

Strobl, C., Malley, J., Tutz, G., & Maxwell, S. E. (2009). An Introduction to Recursive Partitioning: Rationale, Application, and Characteristics of Classification and Regression Trees, Bagging, and Random Forests. *Psychological Methods*, 14(4), 323-348.

Hothorn, T., Hornik, K. and Zeileis, A. (2006). Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics*, 15(3), 651-674.



Stability and longevity in the publication careers of U.S. doctorate recipients¹

Cathelijn J. F. Waaijer^{*}, Benoît Macaluso^{**}, Cassidy R. Sugimoto^{***} and Vincent Larivière^{****}

^{*}cathelijnwaaijer@gmail.com

Centre for Science and Technology Studies, Faculty of Social and Behavioural Sciences, Leiden University, P.O. Box 905, Leiden, 2300 AX (the Netherlands)

^{**}macaluso.benoit@uqam.ca; lariviere.vincent@uqam.ca

Observatoire des sciences et des technologies, Centre interuniversitaire de recherche sur la science et la technologie, Université du Québec à Montréal, P.O. Box 8888, Montréal, H3C 3P8 (Canada)

^{***}sugimoto@indiana.edu

School of Informatics and Computing, Indiana University, 1320 E. 10th Street, Bloomington, Indiana, 47405-3907 (United States)

^{****}vincent.lariviere@umontreal.ca

École de bibliothéconomie et des sciences de l'information, Université de Montréal, 3150, rue Jean-Brillant, Montréal, H3T 1N8 (Canada)

ABSTRACT

Since the 1950s, the number of doctorate recipients has risen dramatically in the United States. In this paper, we investigate whether the longevity of doctorate recipients' publication careers has changed. This is achieved by matching 1951-2010 doctorate recipients in astrophysics, chemistry, economics, genetics and psychology with rare names in the dissertation database ProQuest to their publications in the publication database Web of Science. Our study shows that post-PhD publication career spans have not changed much in most fields, with the share of doctorate recipients who have published for over twenty years having remained stable over time.

INTRODUCTION

The career system in modern day academia is typically pyramidal in structure with relatively few professors at the top and many PhD students at the bottom. Supply typically exceeds demand: more PhD students and postdoctoral researchers indicate they would like to have a university research career than there are positions available (Cyranoski et al. 2011; 'Make the most out of PhDs', 2015; Stephan, 2012). Indeed, studies have shown that opportunities to occupy tenured or tenure track faculty positions in academia have decreased (Schuster & Finkelstein, 2006; Stephan, 2012; Waaijer, 2015).

However, this does not necessarily mean that there are fewer opportunities to do scientific work after the PhD, as "investigative careers" can be found in a number of positions within the scientific workforce. In this paper, we assess the length of the investigative careers of

¹ This work was supported by the Leiden University Fund, grant #5519 for a short research stay in Montréal. A longer version of this paper has been submitted for publication to *PLoS ONE*.

doctorate recipients with rare names from U.S. universities in five fields: astrophysics, chemistry, economics, genetics and psychology, for the period 1951 – 2010.

DATA AND METHODS

The ProQuest Dissertations and Theses database was used as a source of data on U.S. doctorate recipients (ProQuest, 2015). From this database, the names of all 1951-2010 doctorate recipients of U.S. universities in astrophysics, chemistry, economics, genetics and psychology were retrieved. In our study, our goal was to estimate the post-PhD publication careers of these doctorate recipients by measuring when they published scientific papers. However, the attribution of papers can be difficult due to the problem of homonymic names, names shared by one or more individuals (Smalheiser & Torvik, 2009). Therefore, we restricted our sample of doctorate recipients to those with a unique name, an approach also followed by Boyack and Klavans (2008), for example. This was achieved by selecting names of doctorate recipients occurring only once in ProQuest with two or three initials. Furthermore, surnames occurring commonly in the Web of Science were removed (i.e., for each surname the number of distinct combinations of surname and initials was counted, and surnames occurring in 100 or more combinations were removed). Finally, we removed names of researchers with publications in fields outside the PhDs' broader research field, as such publications suggests these names are homonyms.

Unique names in the ProQuest database are not necessarily unique in the much larger WoS database. Therefore, further criteria were imposed on the retrieval of papers from this database. The type of paper was limited to “articles” or “reviews” as we are interested in measuring the research output of doctorate recipients. Papers published between five years before PhD and thirty years thereafter were retrieved. This period was chosen because publications published long before the PhD (e.g., ten years before PhD), or many years after (e.g., sixty years after) are unlikely to be authored by the doctorate recipient. As a further selection criterion a doctorate recipient's first publication must be between five years before and three or five years after PhD graduation; three years for astrophysics, chemistry, and genetics; five years for economics and psychology as PhD graduates in these fields publish their first paper later than in the other three fields. In addition, at least one publication must be in the (narrow) field of their PhD.

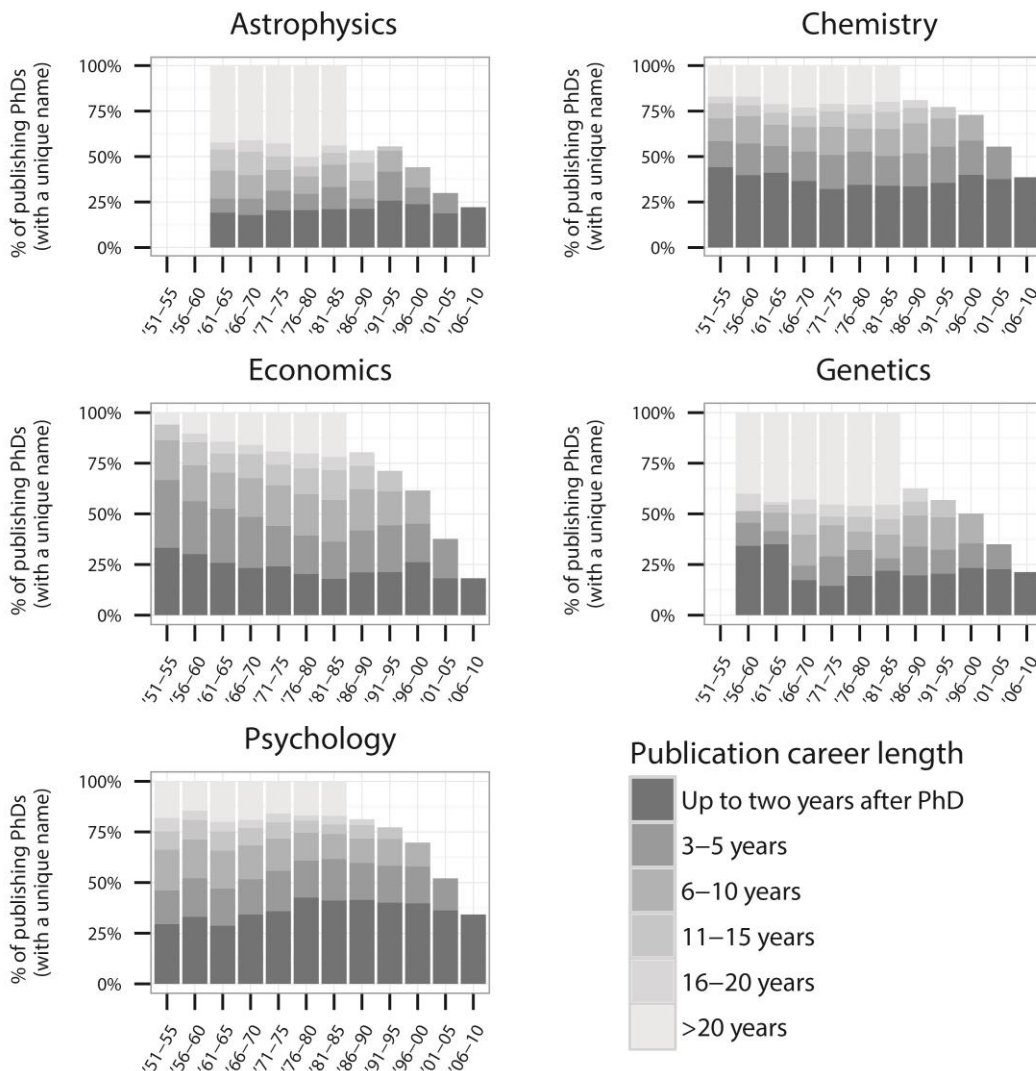
RESULTS

Post-PhD career spans of U.S. doctorate recipients were examined by computing the share of recipients publishing at various career lengths: up to two, three to five, six to ten, eleven to fifteen, sixteen to twenty, and twenty to thirty years since the doctorate. We consider the publication career spans of the 1951-2010 doctorate recipients, with papers published after a long interruption in publication (five years or longer) removed (Fig. 1). In such an investigation, no distinction can yet be made between post-1985 doctorate recipients with a short publication career and those whose publication careers have been interrupted but who will later resume publishing. For example, a scholar receiving a PhD in 1990 could have published their last paper in 1995, which means a career span of five years. However, they could publish a next paper in 2017, which would mean their career span would actually be 27 years. Clearly, it is impossible to predict if this will happen. However, when papers published after an interruption are disregarded, the determination of post-1985 doctorate recipients'

publication career length is possible. Results are only shown when all doctorate recipients in a five-year period have had the opportunity to publish in a given period. An example: for 2001-2005 doctorate recipients, we do not plot the shares of doctorate recipients publishing for 6-10 years after PhD, because this figure also includes people that will continue to publish for a longer period. Hence, the shares do not add up to 100%. Finally, shares were only plotted when number of doctorate recipients with one or more published papers in a five-year period was > 25.

Looking at trends in the spans of publication careers, in astrophysics, chemistry, genetics and psychology, the share of doctorate recipients with long publication careers has remained quite stable, but increased in economics. With respect to the share of brief publication careers, in chemistry, recent decades have seen a slight upward trend after an initial decline. In economics, there was a downward trend. The share of intermediate length careers (6-20 years) decreased in psychology in the late 1980s and the 1990s.

Figure 1. Career length by five-year period and field.



DISCUSSION AND CONCLUSIONS

The large growth in the number of doctoral students compared to a smaller growth in tenure track positions has raised concerns in the scholarly community ('Indentured labour', 2007; 'Make the most out of PhDs', 2015). We investigated whether the changes in academic employment have replaced long investigative careers with more volatile ones for doctorate recipients.

We do not find evidence for changes in academic employment to have led to shorter investigative careers. The span of the publication career has remained stable (in astrophysics, chemistry, genetics, and psychology) or even increased (in economics). Furthermore, in the basic research fields of astrophysics and genetics, long publication careers (of over twenty years) have been the most common career for doctorate recipients from the early 1950s to the early 1980s.

In conclusion, not only are long publication careers common, the shares of more recent doctorate recipients publishing for a short period after the PhD are also stable. Therefore, while employment structures may have changed, the span of research activity by doctorate recipients has not. So in what types of positions do these academics work? Data on academic positions show a large increase in the number of postdoctoral positions (Cantwell & Taylor 2015; Stephan & Ma, 2005). This rise is due to both an increase in the number of recent doctorate recipients taking a first postdoctoral position and to an increase in the time spent in postdoctoral positions (Stephan & Ma, 2005). In addition, the number of non-tenure track staff positions has increased through time (Schuster & Finkelstein, 2006). Our results show researchers have publication careers that are as long as they were before. However, they may be in "holding positions" or on "soft money" (i.e., postdoctoral and non-tenure track positions) for a much longer time. In addition, they may have continued publishing in non-academic employment.

REFERENCES

- Boyack, K.W. & Klavans, R. (2008). Measuring science-technology interaction using rare inventor-author names. *Journal of Informetrics*, 2, 173-182.
- Cantwell, B. & Taylor, B.J. (2015). Rise of the Science and Engineering Postdoctorate and the Restructuring of Academic Research. *Journal of Higher Education*, 86, 667-696.
- Cyranoski, D., Gilbert, N., Ledford, H., Nayar, A. & Yahia, M. (2011). The Phd Factory. *Nature*, 472, 276-279.
- Indentured labour (2007). *Nature*, 448, 839-840.
- Make the most of PhDs (2015). *Nature*, 528, 7.
- ProQuest LLC (2015). *ProQuest Dissertations & Theses Global*. Retrieved July 28, 2015 from <http://www.proquest.com/products-services/pqdtglobal.html>.
- Schuster, J.H. & Finkelstein, M.J. (Eds.). (2006). *The American Faculty: The Restructuring of Academic Work and Careers*. Baltimore: The Johns Hopkins University Press.

Smalheiser, N.R. & Torvik, V.I. (2009). Author Name Disambiguation. *Annual Review of Information Science and Technology*, 43, 287-313.

Stephan, P. (Ed.). (2012). *How Economics Shapes Science*. Cambridge: Harvard University Press.

Stephan, P. & Ma, J. (2005). The increased frequency and duration of the postdoctorate career stage. *American Economic Review*, 95, 71-75.

Waijjer, C.J.F. (2015). The Coming of Age of the Academic Career: Differentiation and Professionalization of German Academic Positions from the 19th Century to the Present. *Minerva*, 53, 43-67.



Developing research career indicators using open data: the RISIS infrastructure¹

Carolina Cañibano* Richard Woolley* Eric Iversen** Sybille Hinze*** Stefan Hornbostel***
Jakob Tesch***

*cacaيسان@upvnet.upv.es; ricwoo@ingenio.upv.es
INGENIO (CSIC-UPV), Universitat Politècnica de València, Valencia, Spain.

**eric.iversen@nifu.no
NIFU, Oslo, Norway

***hinze@dzhw.eu; hornbostel@forschungsinfo.de; tesch@forschungsinfo.de
DZHW (German Centre for Higher Education Research and Science Studies), Berlin, Germany.

ABSTRACT

This paper introduces the research infrastructure for research and innovation policy studies (RISIS) and its ongoing work on the development of indicators for research careers. The paper first describes the rationale for developing an information system on research careers. It then uses an example to demonstrate the possibilities arising from aggregating open data from different datasets within the RISIS platform to create new information and monitoring possibilities with regard to research careers.

INTRODUCTION

The need for developing research careers indicators is driven by three main perceptions: that the science and research systems are changing in ways that have profound impacts on research careers; that systematised knowledge about research careers is lacking; and that transformation in data access opens new avenues for research into scientific research careers.

A very large literature exists that attempts to describe and analyse the changing „contract“ between scientific research and society. Much of this literature has emerged in the past two or three decades. A set of common themes run through this literature:

- there has been a transformation in the organisation of scientific knowledge production (Gibbons et al. 1994, for a survey see Hessels & van Lente 2008);
- public sector organizations involved in the production and utilization of scientific knowledge have experienced transformations in their missions and their expected response to societal expectations (Etzkowitz & Leydesdorff 2000);
- the matrix of authority relations involving governments, funding agencies, research performing organizations and researchers has been in a state of semi-permanent reform or revolution (Whitley, Gläser & Engwall 2010);
- the organization of research has become more bureaucratized and industrialized, with resource allocation and governance moving increasingly to the project level (Walsh & Lee 2015);

¹ This paper was developed as part of the [RISIS project](#) funded through the 7th Framework Program ([DOW RISIS \(313082\) 2013-11-20-1](#)). Funding from RISIS made possible the travel and workshops that allowed the paper's collaborative development. RISIS also directly funded some of the research time devoted to the development of this work.

- the framework conditions of science in a „dynamic steady state“ (Ziman 1994) with increasing competition for finite resources has led to substantive problems with the operation of scientific labour markets (Stephan 2013).

The combination of these changes has apparently had two main effects on research careers. First, an enlarged cohort of researchers clustered at the post-doctoral level has developed in some scientific fields. In many cases these researchers will be employed on consecutive post-doctoral contracts funded by project „soft money“ (Stephan 2014). Second, global emphasis on the circulation of knowledge as a driver of innovation and socio-economic development has contributed to the promotion of scientific mobility through collaboration policy agreements and support mechanisms (Jonkers 2010; OECD 2010). In the case of Europe, promotion of an integrationist policy framework (including the European Research Area (ERA) that modifies national labour market and other rules facilitates researchers moving between Member States (MS) to work in research. These effects are likely to lead to a reduction in the dominance of the single-organization research career form and to an enlarged „tracking task“ for research into research careers (Franzoni, Scellato & Stephan 2012, 2014; Geuna et al. 2015).

Second, the state-of-the-art research on research careers is significantly fragmented and exhibits substantial gaps, including:

- the lack of a comprehensive approach to researching research careers. The complex nature of research careers means that research tends to focus either on institutional determinants or individual choice, but attempts are rarely made to link the two approaches;
- there is a conspicuous lack of attention to research careers outside of academic research careers – and this research generally neglects that university researchers also have significant teaching, administration and other roles;
- comparability between cohort studies of research careers is lacking, not least because there have been very few attempts to construct variables based on universal categories (Kaulisch & Salerno 2005);
- a limited capacity to account for temporality and the evolution of institutions due to the emphasis on cross-sectional data;
- a lack of clarity about the range of research jobs that exist (Miller & Feldman 2014). Some research roles appear to be „invisible“ or appear only as „stepping stones“ to the established career pathway when there is evidence that these roles may be persisting and form significant element of contemporary careers (Miller & Feldman 2014; Stephan 2013);
- a general lack of understanding of mixed and hybrid careers (Lin & Bozeman 2006), although some recent attempts have been made to understand the impacts on researchers“ career aspirations resulting from working in hybrid public/private organisational forms (Garrett-Jones, Turpin & Diment 2013; Gray et al. 2011; Lam 2011).

Third, a growing momentum exists within science and policy communities with regard to organising the practice and communication of science in a more „open“ fashion.

- There is an increasing focus on making research data more openly accessible, both for reuse and to facilitate reproduction studies and validation trials (Crouzier 2015; Moedas 2015).

- The development of significant online architectures for the distributed production, diffusion and use of research data is accelerating in many fields, including the social sciences.

The Research Infrastructure for Research and Innovation Policy Studies ([RISIS](#)) is a project that takes as its primary objective the building of a distributed infrastructure of data relevant for research and innovation dynamics and policies (RISIS nd). It aims to achieve this objective through three complementary activities: opening and developing datasets on issues of continuing relevance to research and innovation research and policy; developing open platforms for building, hosting, curating and developing ad-hoc datasets on research and innovation issues; and developing free and open standards facilitating the interconnection and aggregation of heterogeneous existing datasets (RISIS nd). Current datasets within RISIS cover five dimensions: ERA dynamics; firm innovation dynamics; public sector research; research careers; and a repository on research and innovation policy evaluations (RISIS nd).

RISIS thus provides a single access point to different types of data and existing information relevant to studying research careers. The available data and information includes:

- ETER – database of characteristics of higher education institutions in Europe;
- Leiden Rankings of universities;
- MORE1 and MORE2 – surveys of researcher mobility inside and outside Europe; and
- ProFile – panel data on training conditions and preferred occupational activities of doctoral candidates in German universities (Haus, Kaulisch and Tesch 2015).

The aim of this paper is to provide an initial demonstration of using the RISIS platform to facilitate the aggregation of these datasets to conduct new enquiries into research careers. The open access to multiple data collections has the potential to lead toward the construction of new indicators that can monitor the evolution of systemic change in research careers.

THE ROLE OF MOBILITY IN RESEARCH CAREERS: AN ILLUSTRATIVE EXAMPLE

Policymakers are interested in the phenomenon of mobility in research careers. Researcher mobility is important in the interest of capturing investment in highly skilled human capital, benefiting from knowledge and technology transfer and for research system capacity building (Cañibano & Woolley 2015). From the point of view of individual careers, international job mobility is generally the marker of a critical juncture that reconfigures the contexts and networks in which researchers work and therefore strongly conditions their careers.

A potential descriptive research question is: in what circumstances will researchers move to another country to work? Addressing this question will require information on the personal and institutional push-pull factors that are most important and most commonly aligned in facilitating mobility.

From the individual perspective, empirical research questions can be formulated for mobile researchers:

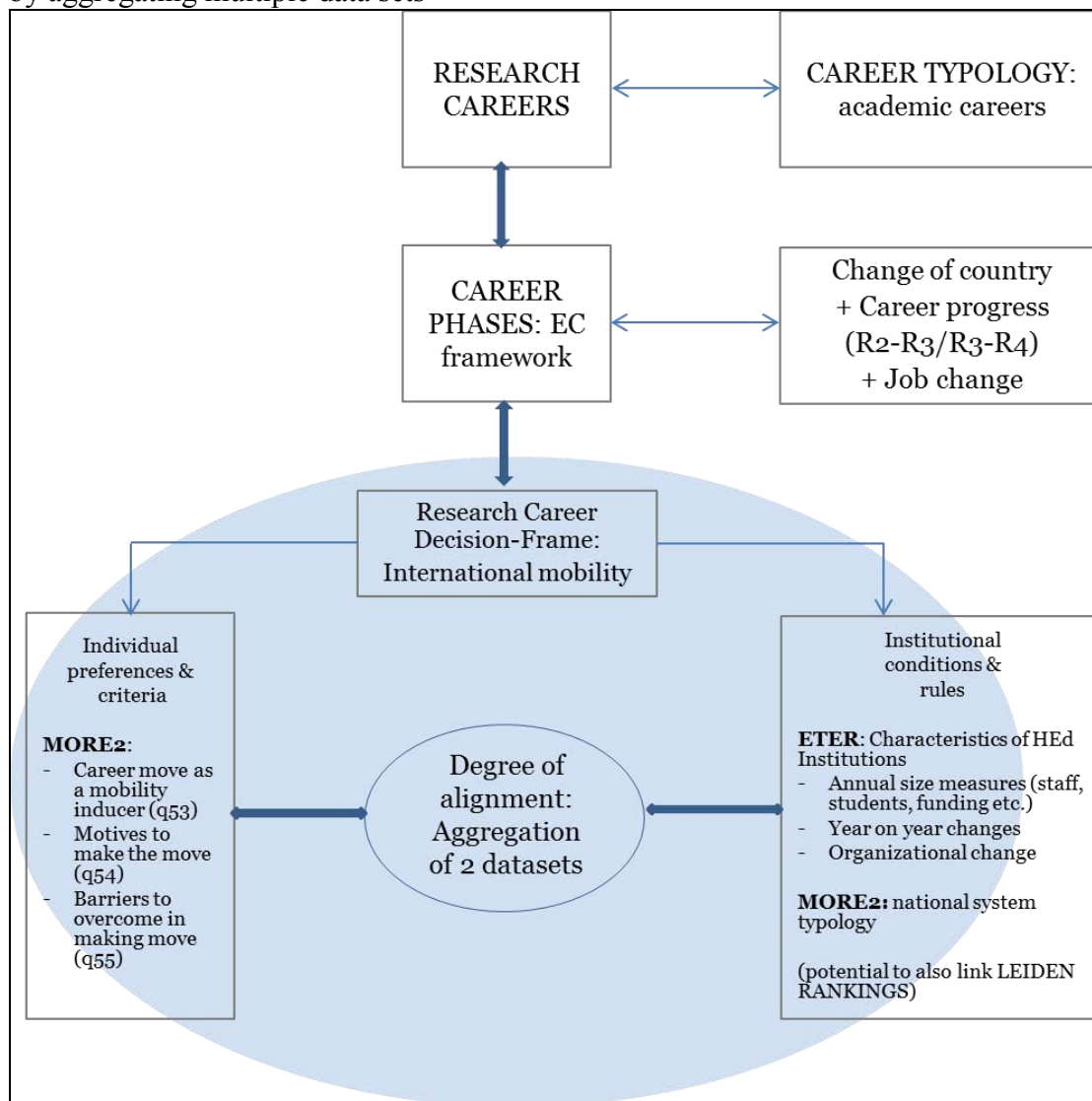
- What motivated you to move to another country to work?
- What obstacles did you encounter in making this move?

From the institutional perspective, empirical research questions can be formulated about the research performing organisations (RPOs) to which researchers move:

- What are the characteristics of RPOs that are „magnets“ for mobile researchers?
- Are the characteristics of „magnet RPOs“ different to those of RPOs overall?
- Are there national system or scientific field effects?

The question „in what circumstances will researchers move to another country to work?“ can be addressed using data available through the RISIS data infrastructure, in particular the MORE1, MORE2 and ETER datasets. These data address academic careers.

Figure 1 Using the RISIS data infrastructure to answer a question about researcher mobility by aggregating multiple data sets

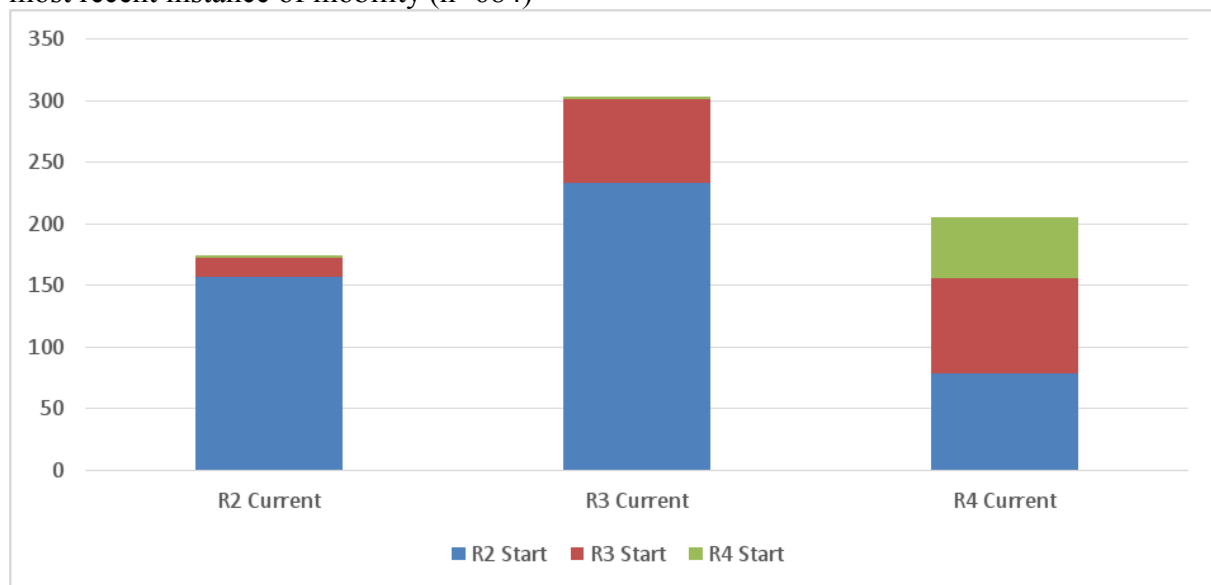


In relation to individual preferences and criteria, the MORE2 (IDEA Consult 2013) dataset can be used to shed light on complex questions associated with decision-frames, including decisions to move to another country. The example descriptive research question focuses on cases where researcher mobility is reported to coincide with critical junctures in career progression. The target population is then researchers who report a change in employer during the past 10 years (MORE2 questionnaire, q50) and/or where the latest episode of mobility coincides with the start-date of their current position (MORE2, q19). Mobile researchers are

also included if they associate their latest mobility spell with some increase in functional status (MORE2, q53). These characteristics exemplify measures of career progression.

Defined in this (fairly restricted) way, the resulting target population provides a lens on decision frames as shaped by institutional factors. The following figure compares the target-population's current level of seniority (in terms of R2-R4) to that reported at the start of their latest mobility spell. The number of respondents in this sub-group (n=684) corresponds to a population of around 6.5 percent (or 82,000) researchers in Europe.

Figure 2 Current career stage of researchers versus their career stage at the beginning of their most recent instance of mobility (n=684)



Source: MORE2 dataset

Researchers who report having functioned as recognized researchers (R2) at the beginning of their most recent geographical move (blue) were generally at an early stage of their careers (and/or at a relatively low level of seniority).

These early stage researchers dominate (72 percent) our target population. We find that a third of these had remained at this level in their current positions, a half had moved to the stage of recognized researcher (R3), while a further 17 percent reported a current status as R4. Similarly, half of those who started their last mobility-stint at R3 had moved to R4 in their current position.

Meanwhile, of those currently in R2 positions (n=175), approximately 10 percent had commenced their latest geographical move in situations equivalent to more advanced career stages, either R4 (n=2) or R3 (n=16). This indicates a career regression which requires further investigation and explanation. In terms of motivations, 80 percent of the target-population reported having been driven by potential for career progression when embarking on a stint abroad as a researcher.

In MORE2 respondents report a great deal of information about their current position, their past positions abroad, and the motivations and barriers experienced along the way. Factors contributing to mobility events such as the types of contracts involved are also reported, in

addition to information about the researcher (age, gender, family situation) that may be correlated with particular career stages. Information is most complete for those respondents who report having international experience that is directly related to their career. Information about the geographical and disciplinary contexts can on this basis be analyzed. For example, one can compare the propensity for researcher mobility in different contexts (geographical, institutional, field) with the researcher's assessment of the home institution (current position).

In terms of the institutional conditions and rules that shape these mobilities observed in MORE2, the outcomes observed can be linked through RISIS to the **ETER** registry of European institutions to determine which types of institutions may be mobility magnets. The ETER data includes annual information (from 2011) about total (academic staff), number of full-professors, revenue, total expenditure, budget, enrolment, graduates, and other measures. Applying year-on-year data from ETER can be used to relative growth or decline according to a given proxy for growth. This information could be used to explore whether annual changes in the university (size, budget, number of students or staff, or performance) correlates with inflows of mobile researchers. The value-added of the RISIS data infrastructure is exploited in this example by linking different research mobility outcomes to certain organisation profiles, providing an overall picture of how individual and institutional factors align in observed mobility patterns. The distribution of poles of attraction for mobile researchers can not just be identified in terms of organisations/countries of destination, but also some additional qualities associated to the research and systemic environment of these mobility magnets can start to be understood.

A further typological variable could also be potentially applied to these linked data. The MORE2 study categorises national academic research systems into a four-category typology depending on a range of institutional and labour market factors. A potential further link could be made through RISIS to a proxy for the overall relative performance, productivity (and hence prestige) of individual universities, through the Leiden Ranking.

DISCUSSION

An important principle of an open science and innovation agenda is the capacity for new users to utilise existing datasets for their self-defined purposes. Opening access to existing datasets can expand the range of questions that these data are used to address. By linking to other datasets the quality of the results that can be generated may also improve. Accessing existing datasets also provides learning opportunities, regarding the types of data collection that might be replicated in a different national or regional context, for example. A case for repeated collection of specialized datasets on research can also more easily be made when the potential exists for these data to benefit users in diverse locations and national systems.

The empirical research questions highlighted in our example demonstrate how the RISIS infrastructure can facilitate such open science and innovation principles. The questions posed are initially addressed through RISIS by aggregating two datasets:

- MORE2 (and MORE1)
 - What motivated you to move to another country to work?
 - What obstacles did you encounter in making this move?
 - and ETER
 - What are the characteristics of RPOs that are „magnets“ for mobile researchers?
 - Are the characteristics of „magnet RPOs“ different to those of RPOs overall?
-

Are there national system or scientific field effects?

However, the potential to link to other datasets, such as ProFile, and information such as the Leiden Rankings, also exists within the RISIS data infrastructure.

The paper demonstrates how the RISIS infrastructure can be used to identify and exploit complementarities between datasets and thus improve our understanding of (changing) research careers. We pointed out several dimensions that can be exploited, including the geographical dimension. An opportunity exists for broad view cross-country surveys to communicate with the more specific national work; points of contact between MORE2 (EU level) and ProFile (Germany) will be explored in future work.

The paper also highlights how RISIS opens possibilities to exploit data complementarities between levels. Researcher-level observations (MORE2) are being enhanced with information about their home institution (ETER) and potentially its relative stature in the higher education/research landscape (Leiden Ranking). In addition, there is a temporal dimension that the infrastructure captures. Some datasets (such as MORE and ProFile) are periodic. Provided internal consistency is maintained, longitudinal information can be collated (MORE2 and the upcoming MORE3) and information accumulated. Indicators with a broad coverage that are fit for monitoring the evolution of research careers over time can thus be the eventual outcome of the RISIS development process. The open data platform approach ensures that use of the available data will be driven by (heterogeneous) research questions and hence open to a diversity of ongoing approaches.

There are of course numerous limitations to these various data and to the interpretations that can be made of the linked data construct at this stage in the development process. The purpose of this methodological paper is to demonstrate the linking of research questions and the RISIS data infrastructure in studying research careers. The paper has outlined the development of a new data sharing and aggregation approach to the development of information about research careers that has the potential to provide a basis for future indicator construction and monitoring development.

REFERENCES AND CITATIONS

Cañibano, C. and R. Woolley. 2015. “Towards a Socio-Economics of the Brain Drain and Distributed Human Capital.” *International Migration* 53(1): 115–30.

Crouzier, T. 2015. *Science Ecosystem 2.0: How will change occur?*, Directorate-General for Research and Innovation, European Commission, Brussels.

Etzkowitz, H. and L. Leydesdorff. 2000. “The Dynamics of Innovation: From National Systems and „Mode 2“ to a Triple Helix of University–industry–government Relations.” *Research Policy* 29(2): 109–23.

Franzoni, Chiara, Giuseppe Scellato, and Paula Stephan. 2012. “Foreign-Born Scientists: Mobility Patterns for 16 Countries.” *Nature biotechnology* 30(12):1250–53.

Franzoni, Chiara, Giuseppe Scellato, and Paula Stephan. 2014. “The Mover’s Advantage: The Superior Performance of Migrant Scientists.” *Economics Letters* 122(1):879–93.

Garrett-Jones, S., Turpin, T. and K. Diment. 2013. “Careers and Organisational Objectives: Managing Competing Interests in Coeprative Research Centres.” Pp. 79–111 in *Cooperative Research Centres and Technical Innovation: Government Policies, Industry Strategies and Organizational Dynamics*, edited by C. Boardman, D. Gray and D. Rivers. New York: Springer.

Geuna, A., Kataishi, R., Toselli, M., Guzmán, M., Lawson, C., Fernandez-Zubieta, A. and B. Barros. 2015. “SiSOB Data Extraction and Codification: A Tool to Analyze Scientific Careers.” *Research Policy* 44(9): 1645–1658.

Gibbons, M. et al. 1994. *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. London: Sage.

Gläser, Jochen and Grit Laudel. 2015. *The Three Careers of an Academic*.

Gray, D., Sundstrom, E., Tornatzky, L. and L. McGowen. 2011. “When Triple Helix Unravels: A Multi-Case Analysis of Failures in Industry–university Cooperative Research Centres.” *Industry and Higher Education* 25(5):333–45.

Hauss, K., Kaulisch, M. and J. Tesch. 2015. “Against All Odds: Determinants of Doctoral Candidates’ Intention to Enter Academia in Germany.” *International Journal for Researcher Development* 6(2): 122-43.

Hessels, L. and H. van Lente. 2008. “Re-Thinking New Knowledge Production: A Literature Review and a Research Agenda.” *Research Policy* 37(4):740–60.

IDEA Consult (2013) MORE2 Final Report. Support for Continued Data Collection and Analysis concerning Mobility Patterns and Career Paths of Researchers. Deliverable 8 prepared for the European Commission. Brussels, August 2013.

Jonkers, K. 2010. *Mobility, Migration and the Chinese Scientific Research System*. Routledge, Oxford.

Kaulisch, M. and C. Salerno. 2005. “Comparing Academic Careers Systems : The Cases of Germany, England and US.” Pp. 1–21 in *30th ASHE conference*. Philadelphia.

Lam, A. 2011. “University-Industry Collaboration: Careers and Knowledge Governance in Hybrid Organisational Space.” *International Journal of Strategic Business Alliances* 2(1/2):135.

Lin, M. and B. Bozeman. 2006. “Researchers’ Industry Experience and Productivity in University-Industry Research Centers: A „Scientific and Technical Human Capital“ Explanation.” *Journal of Technology Transfer* 31(2): 269–90.

Miller, J. and M. Feldman. 2014. “The Sorcerer’s Postdoc Apprentice: Uncertain Funding and Contingent Highly Skilled Labour.” *Cambridge Journal of Regions, Economy and Society* 7: 289–305.

Moedas, C. 2015. Open Innovation, Open Science, Open to the World, speech to European Commission, 22 June, retrieved from <https://ec.europa.eu/research/openscience/index.cfm>

OECD. 2010. “International Mobility.” in *Measuring Innovation: A New Perspective*. OECD Publishing. DOI:10.1787/9789264059474-25-en

RISIS (nd) RISIS: Background on the project, retrieved from <http://risis.eu/about/>

Stephan, P. 2014. “The Economics of the Postdoctoral Position.” Presentation to the National Science Foundation, April, retrieved from <http://sites.gsu.edu/pstephan/economics-of-postdoctoral-position-stephan/>

Walsh, J. and Y-N. Lee. 2015. “The Bureaucratization of Science.” *Research Policy* 44(8):1584–1600.

Whitley, R., Gläser, J. and L. Engwall. 2010. *Reconfiguring Knowledge Production: Changing Authority Relationships in the Sciences and Their Consequences for Intellectual Innovation*. Oxford University Press, Oxford.

Ziman, J. 1994. *Prometheus Bound: Science in a Dynamic “Steady State*. Cambridge University Press, Cambridge.



On the extent of researcher mobility and indicators for mobility

Stina Gerdes Barriere^{*}, Patrik Baard^{*}, Staffan Karlsson^{*}, Carl Jacobsson^{*} and John Tumpane^{**}

^{*} stina.gerdesbarriere@vr.se

Swedish Research Council, Box 1035, SE-101 38 Stockholm (Sweden)

^{**} john.tumpane@formas.se;

Formas (The Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning), Box 1206, SE-111 82 Stockholm (Sweden)

ABSTRACT

Researchers' mobility is significant for the quality of research, although mobility does not necessarily have an intrinsic value. Sweden appears to have a lower degree of mobility when compared to a selection of successful research countries. The Swedish Research Council shall, according to the instruction from the government, promote researcher mobility. This study explores job-mobility within the Swedish academic system based on register studies and draw some conclusions on the degree of mobility in different disciplines, differences between men and women as well as on the development of comprehensive and easy to follow indicators for mobility. The study intends to provide support for recommendations on researcher mobility within the Research Council's overarching goal to promote excellence in research.

BACKGROUND AND PURPOSE OF THE STUDY

Researchers' mobility is significant for the quality of research, although mobility does not necessarily have an intrinsic value. The beneficial effects mobility are gaining access to international contacts and networks, acquiring improved research skills, getting access to infrastructure, etc. While mobility is important to the development of research, and presumably also to the researcher's long-term career progression, there are few measurable short-term effects. (Barriere, 2013) (MORE2 - Higher Education Sector Report, 2013) (Sandström, 2009) (Fernandez-Zubieta, Geuna, & Lawson, 2015) In addition, existing policy instruments and funding models do not necessarily promote mobility.

Sweden has a lower degree of mobility when compared to a selection of successful research countries. (cf. MORE2) Weak career paths and shortcomings in the recruitment process appear to be some of the most important obstacles to mobility, perhaps mainly among researchers who are in the career development stage. For established researchers, the difficulties mainly consist of freeing up time from teaching and administration.

Analysis of mobility needs to take into account and distinguish between the actual mobility and the effects of mobility. We divide this into three different steps:

- the extent of mobility and the patterns of mobility,
- the importance of mobility for the research endeavour
- the effects of mobility on the researcher's career.

The Swedish Research Council shall, according to the instruction from the government, promote researcher mobility. Producing recommendations in matters concerning research

mobility requires studies of several different perspectives. The Swedish Research Council has therefore adopted a long-term perspective in which analyses of the various aspects of researcher mobility gradually contribute to building a knowledge base regarding the prerequisites for and effects of mobility. This paper gives a progress report on this strategy, with focus on indicators for (long-term) geographic mobility within academia. This study does not cover the equally important mobility between sectors of society nor between academic disciplines. In this paper the discussion is limited to mobility that lasts more than three months and with emphasis on long term mobility that includes change of employer. (Vetenskapsrådet, 2015)

Five patterns of geographical mobility can be discerned.

- 1) National mobility – between Higher Education Institutions (HEI:s) within the country
- 2) Temporary outward mobility for shorter periods of time, such as sabbaticals and research visits to longer periods, such as a post-doctoral stays.
- 3) Temporary inward mobility for shorter periods of time, such as sabbaticals and research visits to longer periods, such as a post-doctoral stays.
- 4) Permanent/long-term inward mobility researchers moving to Sweden more or less permanently
- 5) Permanent/long-term outward mobility researchers moving from Sweden more or less permanently

Analyses that inform policy decisions need to distinguish between these different movement patterns as well as prioritize and pinpoint where action is needed or would have the greatest effect. Analyses should also take into account what stage of a researcher's career the mobility takes place.

METHODS

The study covers employees at Swedish institutions of higher education (HEI), holding a doctoral degree and holding teaching and/or research positions (academic staff) and were employed at a Swedish HEI in 2014. This study covers only mobility between the around 40 Swedish HEIs and not between academia and other societal sectors.

The study is divided into two different sub-studies where the first study analyses where researchers and teachers working in higher education gained their doctoral degree in relation to the HEI where they are employed. (Vetenskapsrådet, 2016) Divided by subject area, employment category and HEI type, analysis has been made of the proportion of researchers and teachers holding a doctoral degree from the same HEI at which they are employed, from another Swedish HEI or from a foreign HEI. The analysis is based on register data from Statistics Sweden. The study does not cover shorter employments, since the register is based on the employment in October each year.

The first study uses the term *internal recruitment* for the recruitment of those holding a doctoral degree from their own HEI. The term *national recruitment* is used for the recruitment of those holding a doctoral degree from another Swedish HEI, while *international recruitment* is used for the recruitment of those holding a foreign doctoral degree, but which were employed by a Swedish HEI on the date on which the statistics were compiled. It should be noted that the study does not take into account any intervening employment at other HEIs or

abroad. This study divides the HEIs into different types: large universities, smaller universities and university colleges. Recruitment patterns for university colleges of fine arts and other higher education providers are not covered by this study.

In the second sub-study total employment mobility between different HEI:s for academic staff has been analysed, taking into account each change of academic employer. The study covers change of employers during the time period 2001-2014. The mobility in terms of the number of different HEIs a person had been employed at was studied with three different delimitations:

1. After doctorate (the HEI awarding the doctorate is not counted)
2. HEI of doctorate is included, only doctorates awarded in Sweden is included (doctorates awarded after approx. 1970 are included)
3. HEI of doctorate is included, all doctorates are included, foreign doctorates automatically counted as two HEIs.

RESULTS

Degree of internal recruitment

Just over 50 per cent of academic staff is, by this definition, internally recruited. Internal recruitment in Swedish higher education has declined somewhat during the period studied, 1997-2014.

Variation between disciplines

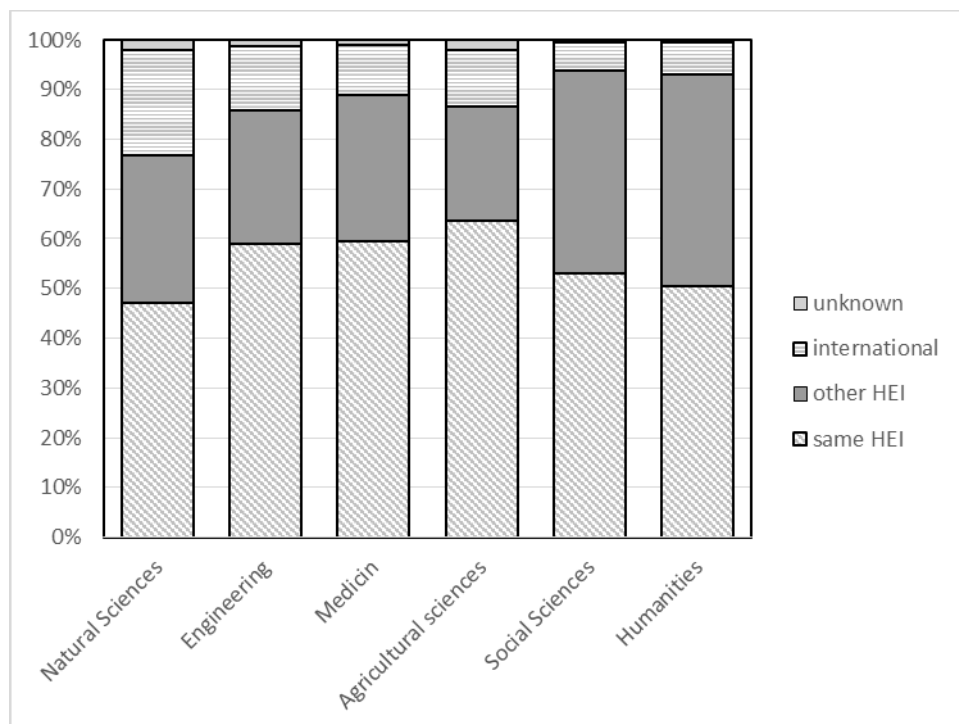
In the natural sciences, internal recruitment is lowest, while the recruitment of those holding a foreign doctoral degree is highest. Medicine and health sciences, and engineering sciences have the next highest level of internal recruitment, around 60 per cent.

About one third of the PhD employees in medicine, engineering and natural sciences gained their doctoral degree from another Swedish HEI. The difference between the disciplines arises primarily from different degrees of international recruitment.

The recruitment profiles for the humanities and social sciences are very similar. In both these subject areas the proportion holding a foreign doctoral degree is just under 7 per cent, while the proportion of employees holding a doctoral degree from another Swedish HEI is just over 40 per cent.

There is some variation between different subject groups within subject areas. Biotechnology, which is studied within four different subject areas – natural sciences, agricultural sciences, medicine and engineering sciences – stands out consistently with a higher level of national and international recruitment when compared to other subjects. Within the social sciences and humanities, the level is fairly even, with the exception of law, which stands out with a higher level of internal recruitment than other subjects. Natural sciences also exhibit few variations between the various subject groups, while the variation is much higher in engineering sciences and in medicine and health sciences. In medicine and health sciences, clinical medicine stands out with a level of internal recruitment of over 70 per cent.

Figure 1: Degree of internal recruitment (Ph.D. from same HEI as current employment), national recruitment (Ph.D. from other Swedish HEI) and international recruitment (Ph.D. from other country) for different scientific disciplines. Data for 2014 from Statistics Sweden.



Differences between men and women

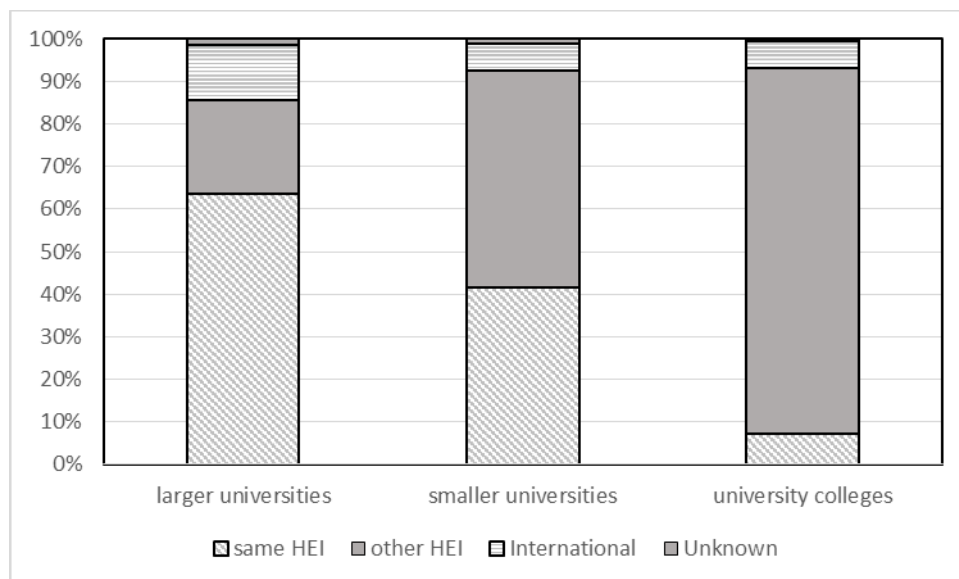
There is a minor but systematic difference between men and women, where women are recruited internally to a somewhat greater extent than men across all scientific disciplines. This difference is found both within staff categories and within subject areas. The difference is accounted for by the fact that a somewhat higher proportion of men than women are recruited internationally.

Variation between HEI types

The large universities have the highest level of internal recruitment, which is particularly prominent in the staff categories of professor and senior lecturer. Recruitment patterns for the smaller universities reflect the fact that most of them have relatively recently (between 1997 and 2005) been granted general degree-awarding powers for postgraduate education. The smaller universities have a lower level of internal recruitment for professors than larger universities, and a relatively high and increasing level of internal recruitment for senior lecturers, and for those in postdoctoral and career-development positions.

The university colleges have a lower level of internal recruitment for all staff categories compared to both large and smaller universities, due to the fact that they do not have general powers to award doctoral degrees. Both the smaller universities and the university colleges also exhibit a lower level of international recruitment compared to the large universities.

Figure 2 Degree of internal recruitment (Ph.D from same HEI as current employment), national recruitment (Ph.D: from other Swedish HEI) and international recruitment (Ph.D. from other county) for different types of HEI:s. Data for 2014 from Statistics Sweden.



Internal recruitment at the large HEIs

The large universities comprise almost 80 per cent of the population investigated. In many cases, they have also existed a significantly longer time period than other HEIs. Due to having existed for a short time period, smaller HEIs often exhibit a greater extent of non-internal recruitment. This is the case for obvious reasons, as several of them do not have, or have not had the right to award post-graduate degrees for very long. However, the data also shows that the degree of internal recruitment increases over time for smaller HEIs. This being the case, it is motivated to take a more focused look solely on large HEIs, since they account for the greatest share of the academic population.

On the whole, the large universities show a great tendency towards recruiting their own PhD's. This is particularly evident for senior lecturers and professors, but is also seen in early career stages such as career-development positions. Career-development and postdoctoral positions more closely reflect the current recruitment climate since these positions can only be held for four and two years, respectively.

For all subject areas combined, just over 60 per cent of professors at the large HEIs have been recruited internally. The highest internal recruitment of professors at the large HEIs is found in medicine and health sciences, where nearly 70 per cent of professors are employed at the same HEI at which they gained their doctoral degree. At the same time, 7 per cent of the professors in medicine hold a foreign doctoral degree.

The senior lecturer category at the large universities exhibits by far the highest level of internal recruitment, with more than 70 per cent holding a doctoral degree from the HEI at which they work. The proportion of internally recruited senior lecturers varies from 60 per cent in natural sciences to over 80 per cent in medicine and health sciences.

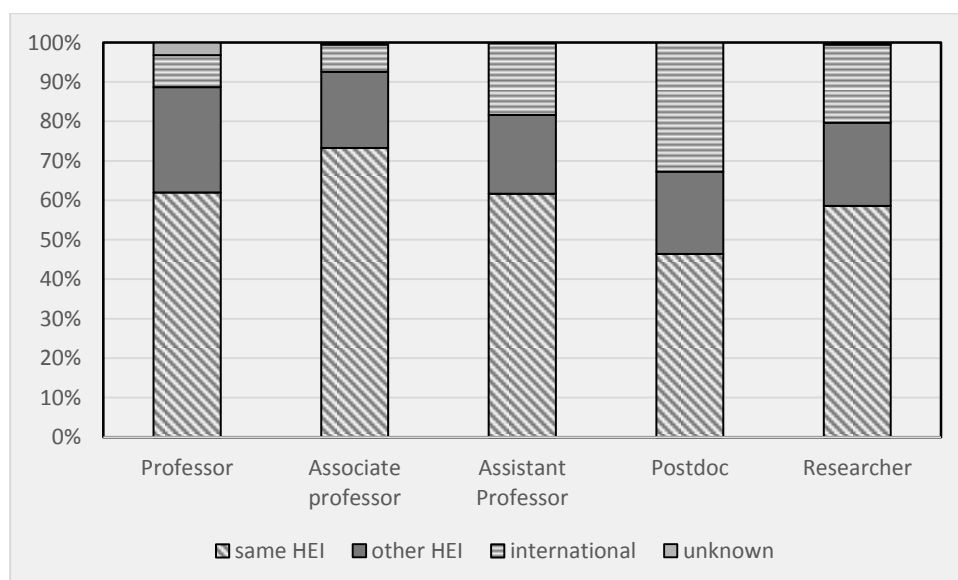
Over 60 per cent of those in career-development positions had obtained their postgraduate education from the same HEI. This is surprising, given that an ambition with such positions is to provide young researchers a clear career path, and thus to attract the most promising

researchers. In natural sciences, a little over half of those in career-development positions were recruited internally, while about two thirds were recruited internally in medicine and health and in engineering. That the proportion of those recruited internally is the same for those in career-development positions as for professors is remarkable, especially since these positions are intended, among other things, to be an entry into a continued career in higher education.

Postdocs at the large universities exhibit the highest level of international recruitment and the lowest level of internal recruitment. In natural sciences, only one third of postdocs holds a doctoral degree from their own HEI, and nearly 45 per cent hold a foreign doctoral degree. In other subject areas, however, internal recruitment is considerable, even for postdocs with nearly half being recruited internally. However, the data should be interpreted with some caution as there is a relatively large group of postdocs who have not applied for the formal conferral of their doctoral degree. It is primarily the proportion of nationally recruited postdocs that is lower compared to other staff positions, while the proportion of postdocs in engineering and the social sciences recruited abroad is higher than other positions. The lowest proportion of postdocs holding a foreign doctoral degree is found in medicine and health sciences.

The level of internal recruitment in the category of researchers at the large universities is over 60 per cent in all subject areas except natural sciences, where just under half hold a doctoral degree from the HEI of current employment, while those recruited abroad vary from just under 30 per cent in natural sciences, in descending order, engineering sciences, medicine and the social sciences, down to around 10 per cent in the humanities.

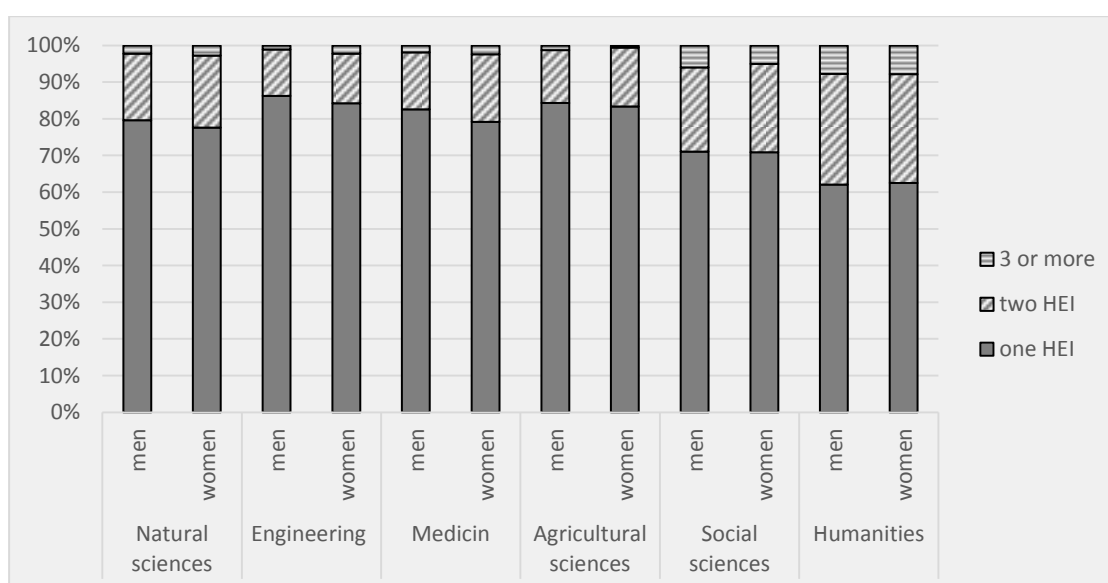
Figure 3 Degree of internal recruitment (Ph.D from same HEI as current employment), national recruitment (Ph.D. from other HEI) and international recruitment (Ph.D. from other county) for different positions at large universities. Data for 2014 from Statistics Sweden.



*Job mobility within academia (Forskarkarriärutredningen, 2016)***Mobility after doctorate**

When looking at mobility for academic staff after doctorate we found that 77 per cent had been employed at only one HEI. Mobility was slightly higher among women compared to men, 78 per cent and 75 per cent respectively. We found some variation between research fields, with highest mobility in the arts and humanities, with 62 per cent non-mobile and lowest mobility in engineering with as much as 86 per cent non-mobile. The difference in mobility between the sexes is small but the higher mobility among women is consistent across all research areas.

Figure 4: Mobility after doctorate (data from 2001-2014), number employers within academia. Data: Statistics Sweden.

**Mobility for Swedish doctorates, including doctorate-awarding HEI**

When the Ph.D.-awarding university is included in the analysis of academic staff with a Swedish doctorate 61 per cent were employed at the same university as the Ph.D.-awarding university and had not changed employer in between. There were no differences between men and women.

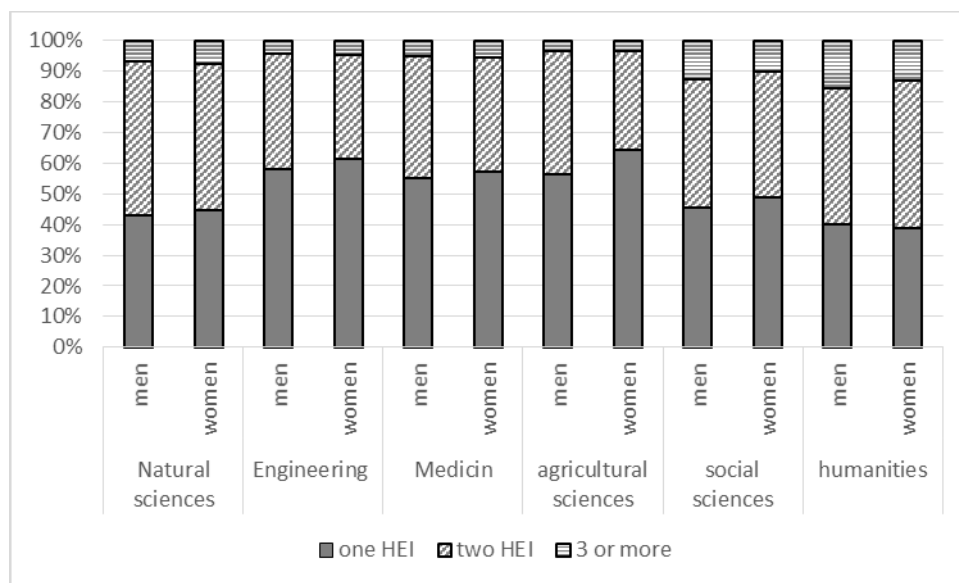
Mobility is, according to this measure, highest in the arts and humanities, with 45 per cent that had been mobile at least once. A little more than 40 per cent had changed HEI once and a little more than ten per cent had changed universities two or more times. Mobility is lowest in engineering and agricultural sciences with about 70 per cent employed at the Ph.D.-awarding university.

Total mobility

When all doctorates are included in the analysis – whether the Ph.D.-awarding institution is Swedish or foreign – total mobility (defined as the change of institution either directly after doctorate or later in their careers) increases the degree of mobility. In average, just half, 50 percent, of the university's academic staff have changed university according to this measure.

The degree of mobility is, on average, the same for men as for women (49 and 51 percent). There is some variation between fields of science: mobility is highest in the humanities, followed by the natural sciences and lowest in agricultural sciences. Within the different disciplines, mobility for women are slightly lower for women than men, in all disciplines except humanities.

Figure 5: Total mobility (job mobility between 2001-2014 and relative to doctorate-awarding HEI) for academic staff 2014. Data source: Statistics Sweden.



DISCUSSION AND CONCLUSIONS

From the first study on internal recruitment we find that about 12 per cent of the academic staff hold a doctorate awarded from a foreign university, which means that they are automatically designated as “mobile” according to the measure used here. A further 33 per cent hold a Swedish doctorate but are not employed at the same university that awarded their doctorate. The remaining 54 per cent are employed at the same university as the one that awarded their doctorate.

The interesting question is whether they were mobile in the intervening period. From the study of the whole population we concluded that 50 per cent of the academic staff had not been mobile, according to the definitions made in the methods section. The most common situation for mobile staff is to change university only once, while less than ten per cent change university twice or more. These results, together with the fact that the degree of internal recruitment is quite high for positions at all career levels. Combining results the two sub-studies suggests that only four per cent of the academic staff had been mobile and returned to the alma mater. The most frequent time period to change academic employer is either directly after the doctorate or between two to five years after Ph.D. These two periods may be intertwined since other studies (Barriere, 2013) show that an international postdoctoral stay is common in the natural sciences, engineering and medicine, although this needs to be verified by further studies.

It is commonly stated that women are less internationally mobile than men, and indeed the difference between men and women in Sweden was large in the MORE-studies (MORE2 -

Higher Education Sector Report, 2013). Our studies indicate, however, that the overall job mobility is as high for women as for men among academic staff, although women may be nationally mobile to a somewhat larger extent while men are more internationally mobile.

The conclusion from these studies is that the mobility between the Ph.D.-awarding university and HEI of (current) employment is a good indicator, although with a slight underestimation, of mobility. This indicator is also easier to obtain than the full analysis of all job-changes. By doing the analysis on type of position in the career system: eg. post-doctoral/researcher/assistant professor/associate professor/full professor, a good understanding of the dynamics in the research system may be obtained. In the case of Sweden, the analysis showed for instance that although the mobility was quite high among post-docs (two years after Ph.D.) The mobility was significantly lower already at the assistant professor stage, giving food for thought about the functionality of the Swedish academic career system.

REFERENCES

- Barriere, S. G. (2013). *PM om Forskarmobilitet, dnr 354-2013-91*. Vetenskapsrådet.
- Fernandez-Zubieta, A., Geuna, A., & Lawson, C. (2015). Productivity pay-offs from academic mobility: should I stay or should I go? *Industrial and Corporate Change*, 25(1).
- Forskarkarriäretredningen. (2016). *SOU 2016:29*. Statens offentliga utredningar.
- (2013). *MORE2 - Higher Education Sector Report*. European Commission.
- Sandström, U. (June 2009). Combining Curriculum Vitae and Bibliometric Analysis: Mobility, Gender and Research Performance. *Research Evaluation*, ss. 135-142.
- Science Europe; SciVal Analytics (Elsevier). (2013). *Comparative benchmarking of European and US research collaboration and researcher mobility*. Science Europe & Elsevier.
- Vetenskapsrådet. (2015). *Forskningens framtid! Svenska forskares mobilitet - en kunskapsöversikt*. Vetenskapsrådet.
- Vetenskapsrådet. (2016). *REKRYTERING AV FORSKARE OCH LÄRARE MED DOKTORSEXAMEN VID SVENSKA LÄROSÄTEN - EN STRUKTURELL ANALYS AV LÄROSÄTENAS REKRYTERING -INTERNT, NATIONELLT OCH INTERNATIONELLT*. Vetenskapsrådet.



Progress on mobility and instability of research personnel in Japan: scientometrics on a job-posting database for monitoring the academic job market

Hiroataka Kawashima ^{*1} and Yasuhiro Yamashita ^{**1}

* kwsahr@nistep.go.jp

2nd Theory-Oriented Research Group, National Institute of Science and Technology Policy (NISTEP), Ministry of Education, Culture, Sports, Science and Technology (MEXT), 16th Floor, Central Government Building No. 7 East Wing, 3-2-2 Kasumigaseki, Chiyoda-ku, Tokyo 100-0013 (Japan)

** yasuhiro.yamashita@jst.go.jp

Office of Information Analysis, Department of Information Planning, Japan Science and Technology Agency, 5-3, Yonbancho, Chiyoda-ku, Tokyo, 102-8666 (Japan)

ABSTRACT

This study has two purposes. The first purpose is to extract statistics from a database of job-posting cards, previously little-used as a data source, to assess the academic job market. The second purpose is to connect statistics on the academic job market with monitoring of indicators of policy progress related to the mobility and instability of research personnel. The data source used in this study is a job-posting database named JREC-IN Portal, which is the de facto standard for academic job seeking in Japan. The present results show a growing proportion of fixed-term researchers in the Japanese academic job market and that job information is increasingly diverse.

INTRODUCTION

Academic researchers seek jobs as individual players in the academic market. The condition of the academic job market is sensitive to science and technology (S&T) policy. Therefore, measuring the academic job market is directly connected to S&T indicators. Academic researchers seeking jobs monitor and search websites that host job-posting databases, perhaps on a daily basis. However, few previous studies have analyzed academic job-posting databases (Gourova and Sanopoulos 2010).

In the case of Japan, most academic researchers seeking jobs use a web service named JREC-IN Portal managed by a public institution. Previous studies have discussed academic human resources based on analysis of the JREC-IN Portal database. Miura and Sato (2007) aggregated data on employer types from job-posting cards and clarified that more than half of all job postings were made by national universities. They concluded that this means the ratio of public recruitment is higher at national universities than private universities. Kobayashi (2008) surveyed job-posting information for physical education teachers and sports scientists and estimated the ratio of job requirement information. Hashimoto et al. (2012) focused on the employment situation of doctoral students in humanities and social sciences and found an increase in the ratio of limited-term employees, especially in non-traditional new research domains. However, the implications of these previous studies are very limited because the

¹ Each of these two authors contributed equally.

data on which they are based are limited both temporally (less than one year of data) and informationally (data items obtainable by manual web scraping).

Here, we analyze the whole of this job-posting database to understand the academic job market in Japan. Japan is an optimal case for the application of bibliometrics to job-posting databases to monitor the academic job market because JREC-IN Portal is established as the de facto standard for academic job seeking, and so covers a sufficient ratio of this market to allow a basis for confident conclusions. Miura and Sato (2007) compared job-posting cards published on JREC-IN Portal with those published on individual university websites over one year. They estimate that 99.8% of job cards published on university websites were also published on JREC-IN Portal.

POLICY RELEVANCE

Japan's 5th Science and Technology Basic Plan (S&T Basic Plan) was decided by the cabinet in January 2016. In Japan the Science and Technology Basic Law mandates that the S&T Basic Plan be determined at five-year intervals. The period of the 5th S&T Basic Plan is from the 2016 to 2020 fiscal years. One significant feature of the 5th S&T Basic Plan is "real-time monitoring". This means that government and policy-related institutions must make and measure indicators for progress management during, not after, the period of the plan. Another significant feature is the mention of job instability in the academic job market. In fact, present Japanese S&T human resource policy is suffering from a trade-off between mobility and instability. This study also aims to contribute "real-time monitoring" of policy progress and the effects of policy on the mobility and instability of research personnel.

DATA SOURCES

The study data comprise all job-posting cards released by universities between 2002 and 2015 on JREC-IN Portal. JREC-IN Portal was launched in October 2001 as an online service for academic job seekers. Table 1 shows the total numbers of job postings by type of organization in 2015 (Calendar year). Since we focus on the recruiting trend of Japanese universities, other non-university institutions are neglected in the present analysis. The targets of the present study are shown in boldface in Table 1. In this study, individual job-posting cards are assigned to the calendar year in which they are published on the web service. Table 1 and our results below also include all types of jobs and positions.

Table 1 Total number of job postings by each type of organization in 2015

Type of organization	Numbers of job postings in 2015
National University	5,108
Prefectural and Municipal University	1,252
Private University	6,647
Two-year College	668
Colleges of Technology	292
Specialized training college	377
Other educational institute	232
Inter-University Research Institute Corporation, Incorporated	2,180

administrative agency and National Institute.	
Public research organization and Local Government	238
Governmental Corporation and Authorized Corporation	8
Public Interest Incorporated Associations	190
Private Company	439
Organization Abroad	57
International Organization	39
The Others	113
Total	17,840

RESULTS AND DISCUSSION

Figure 1 shows the number of job postings for tenured versus fixed-term employees, and Figure 2 shows the number of job postings for full-time versus part-time employees. These two time series clearly show the rapid change of the Japanese academic job market. Although in this market the size of the "pie" has grown, the composition of the filling has changed such that it no longer nearly exclusively comprises full-time and tenured employment, but rather includes a growing proportion of fixed-term employment. Table 2 shows the rate and number of tenure job postings by university type. "Research University 11" (RU11) refers to a group of nine large national universities and two large private universities, all of which have large-scale research activities. The group labelled "Private Universities" is separated into two groups using the threshold of 100 job postings totally on JREC-IN Portal. The "-2" in the group "Large Private Universities - 2" denotes the exclusion of the two universities included in the RU11 group. Table 2 indicates a drop-off in the rate of tenure, especially for the RU11 group, which collectively represents the most important platform for academic research in Japan.

Figure 3 shows vocabulary on job information. The number of unique nouns used in each job description is counted. According to Figure 3, the contents of job postings have become increasingly diversified. A possible reason for this increasing complexity, especially from 2002–2005 to 2007–2010, is the transformation of Japanese national universities into independent administrative entities in fiscal year 2004. For example, many names of universities with legal status first appeared in 2007–2010, as did terms such as "community health nursing", "iPS", "science communicator", "portfolio", "tenure track" and "female researcher". After this transformation, each national university was compelled to characterize itself or seek unique points of difference from other institutions. Therefore the names of departments or descriptions of teaching and research missions came to require an increasingly extensive vocabulary.

Figure 4 aims to show the effect of policy on diversity. Figure 4a shows the number of job postings for which the job information contains the phrase "tenure track" and the breakdown by type of organization (university). In the Japanese context, the tenure-track system is a politically introduced tool designed to enable young researchers to make the transition from fixed-term to tenured employment. Figure 4b shows the number of job postings for which the job information contains the phrase "gender equality" and the breakdown by type of organization (university). Gender equality is also politically promoted in Japan. Remarkably, RU11 and other national universities represent a large share of such job postings and their

share is growing rapidly. This means they are more politically sensitive than prefectural and municipal universities or private universities.

Figure 5 aims to show the difference in circumstances among research fields (Figure 5a) and job types (Figure 5b). The job postings in Figure 5a and 5b are limited to full-time ones. Figure 5a quantitatively displays the cruel fact, suspected qualitatively, that especially in the Biology / Biological domain, most job postings are for fixed-term positions. This fact confirms the existence of a structure whereby a small number of principal investigators employ and lead a larger number of team members to complete terminable projects with fixed budgets. Figure 5b shows another aspect of this same situation. In Japanese universities, almost all researcher/postdoc positions without teaching load are based on specific terminable projects. So institutionally there are very few or no tenured researcher/postdoc positions. However, Figure 5b indicates that most job postings suited to young researchers in all fields are for fixed-term employment. This trend is strengthening as is shown in Figure 1.

CONCLUSION

Here, we show both progress on mobility and employment instability affecting research personnel in Japan by performing scientometric analysis of a job-posting database for monitoring the academic job market. We tried to extract statistics from a database of job-posting cards, previously little-used as a data source for scientometrics or monitoring of policy effects. The extracted statistics indicate the circumstances of the academic job market in Japan and reveal numerous alarming phenomenon related to the mobility and instability of research personnel. The authors hope this study can contribute to policy making and lead to future improvement of job circumstances for academic research personnel.

Figure 1: The number of job postings for tenure/fixed-term positions

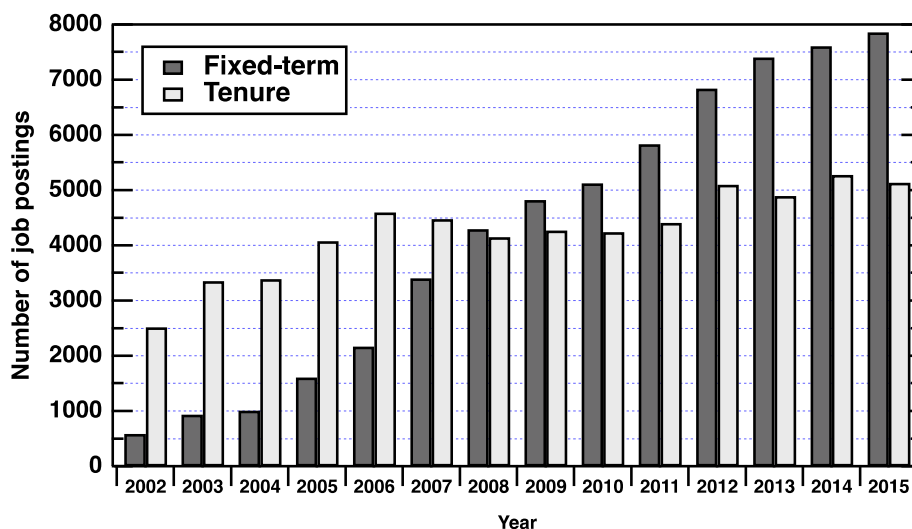


Figure 2: The number of job postings for full-time/part-time positions

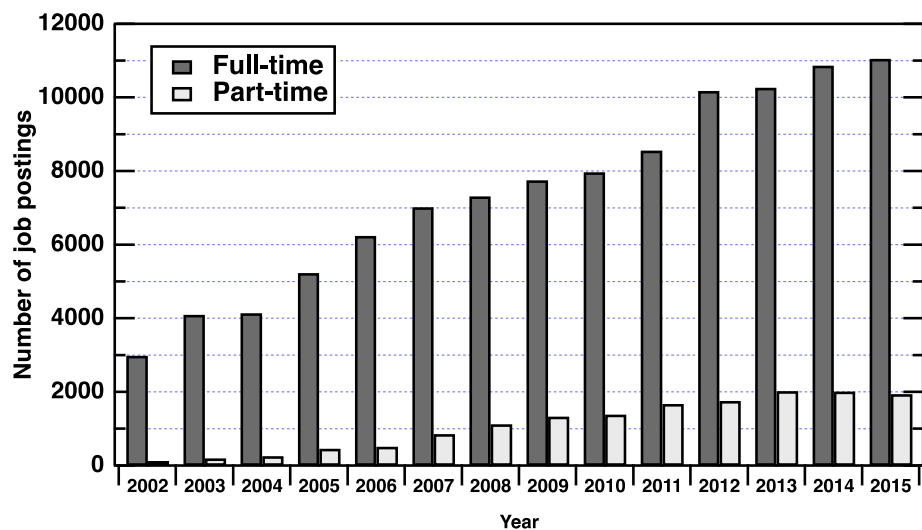


Table 2: The rate (number) of tenure job postings by type of university

	2002-2005	2007-2010	2012-2015
All Universities	76% (13358)	49% (17138)	41% (20419)
Research University 11	65% (2257)	34% (2333)	28% (2576)
Other National Universities	80% (4615)	54% (5426)	46% (5635)
Prefectural and Municipal Universities	80% (1702)	54% (2227)	45% (2175)
Large Private Universities – 2	80% (3443)	52% (5145)	42% (6861)
Small Private Universities	75% (1341)	51% (1903)	42% (3172)

Figure 3: Vocabulary used in job information: the number of unique nouns

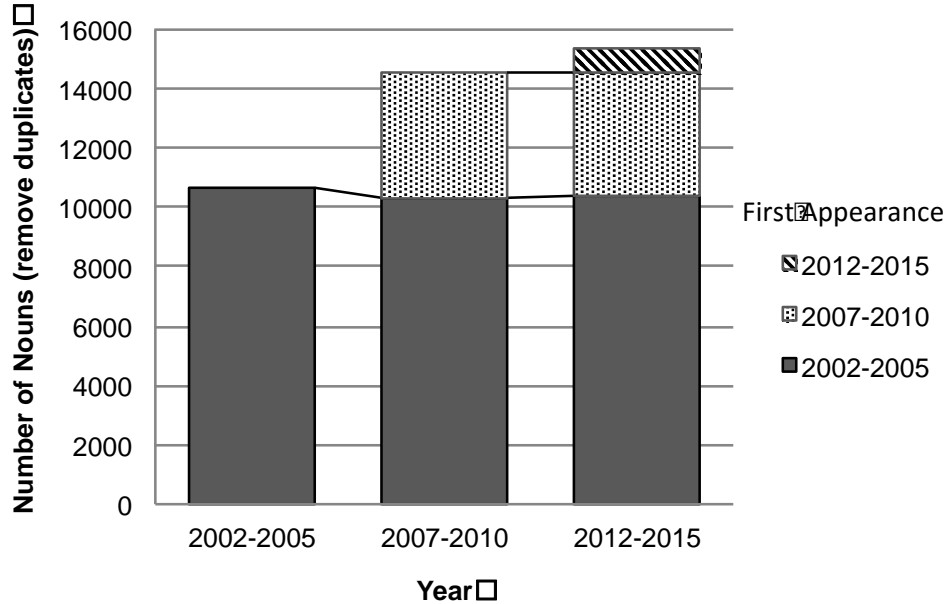
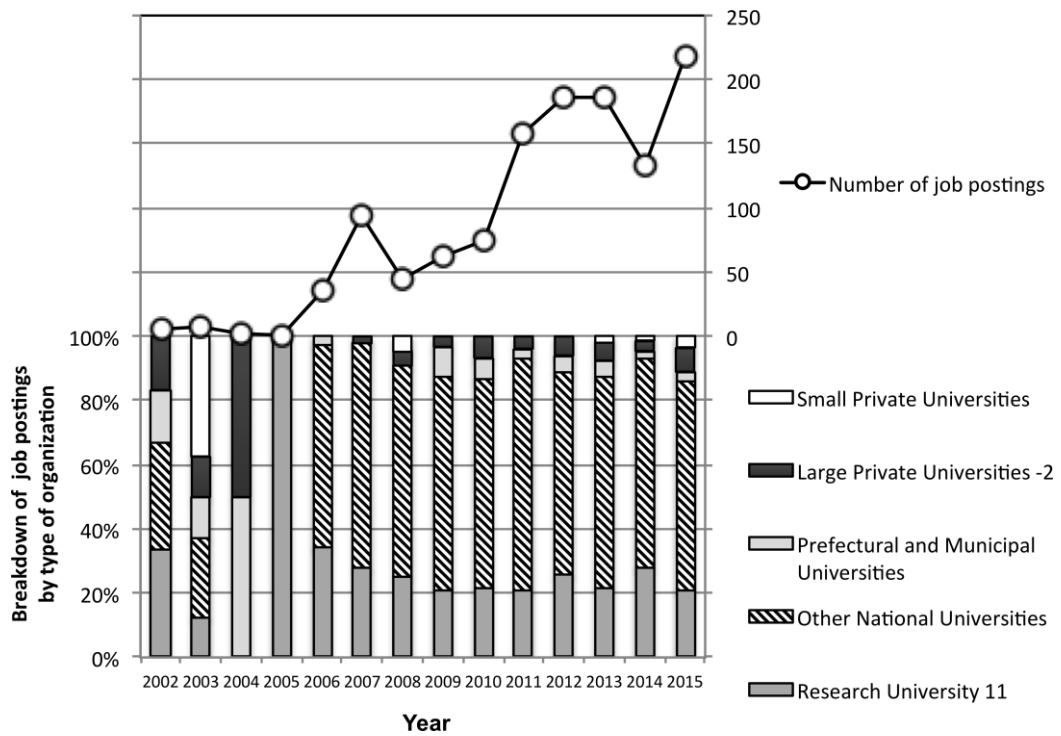


Figure 4: Effect of policy on tenure track and gender equality

a. The number of tenure track-related job postings



b. The number of gender equality-related job postings

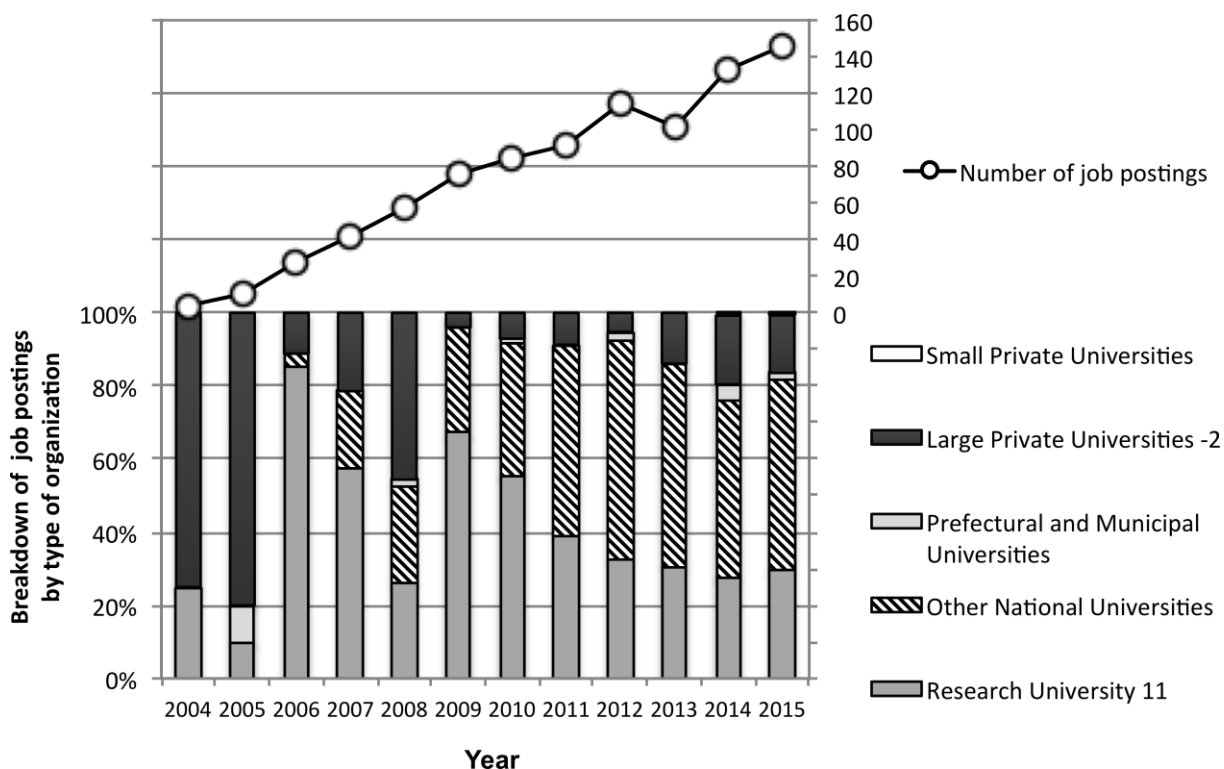
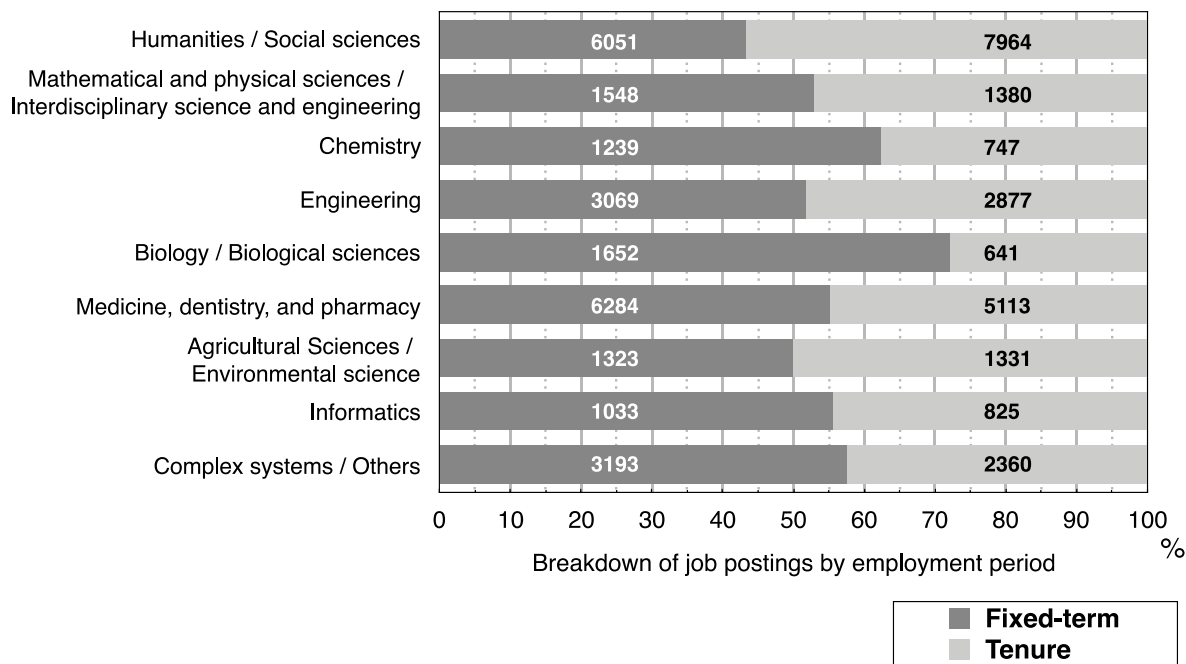
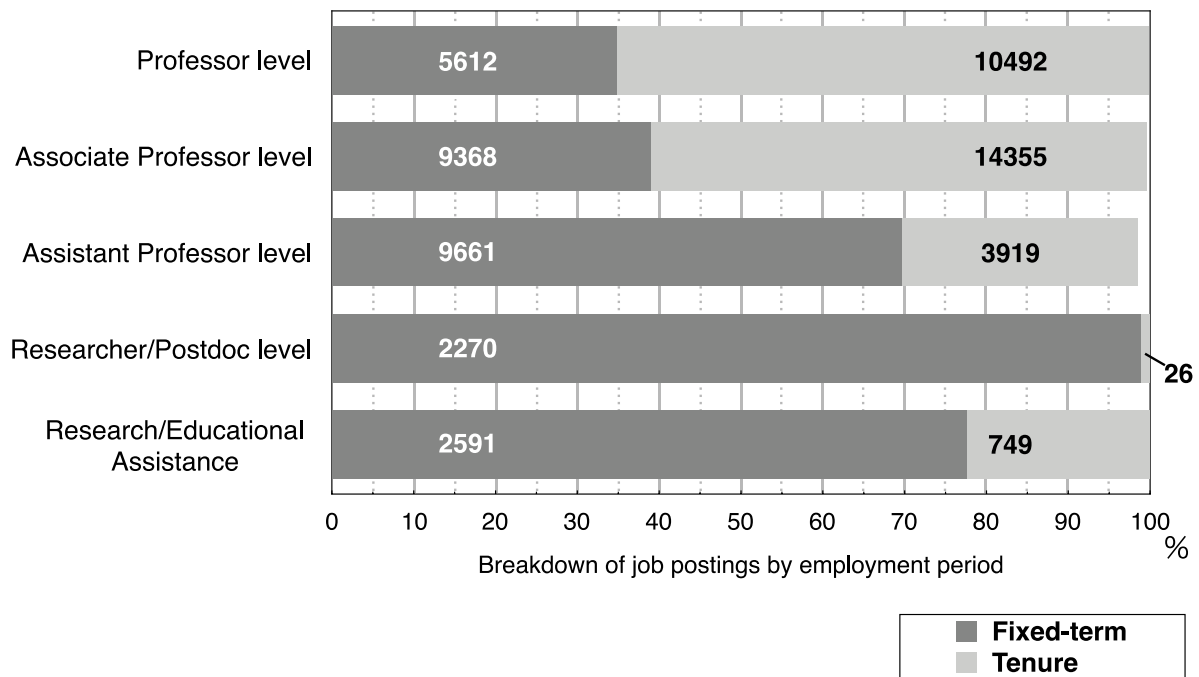


Figure 5: The number of full-time job postings published between 2012 and 2015

a. Breakdown by research field



b. Breakdown by job type



REFERENCES

Gourova, E. & Sanopoulos D. (2010). Knowledge transfer and mobility: EURAXESS role in europe. *International Conference on Engineering Education and International Conference on Education and Educational Technologies – Proceedings*, 349-354.

Hashimoto, K., Saito T., Kato Y. & Chida K. (2012). Employment Situation of Doctoral Students in Humanities & Social Sciences and their Employability in the Academic Market. *Bulletin of the Graduate School of Education, the University of Tokyo* (published in Japanese), 52, 61-86.

Kobayashi, K. (2008). An analysis on recruitments of physical education teachers or sports scientists during 2006 academic year in Japan. *Japan Journal for Health, Physical Education, Recreation, and Dance in Universities* (published in Japanese), 5, 53-60.

Miura, Y. & Sato S. (2007). Recruitment of Researchers by Universities and Public Research Institutions. *Research Material published by National Institute of Science and Technology Policy* (published in Japanese), 133.



National and international scientific elites: an analysis of Chinese scholars¹

Fei Shu^{*}, Vincent Larivière^{**} and Charles-Antoine Julien^{*}

^{*}fei.shu@mail.mcgill.ca; charles.julien@mcgill.ca

School of Information Studies, McGill University, 3661 Peel, Montreal (Canada)

^{**}vincent.lariviere@umontreal.ca

École de bibliothéconomie et des sciences de l'information, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal (Canada)

ABSTRACT

The purpose of this study is to compare the WoS with a national Chinese bibliometric database at the level of individual authors and measure the extent of the overlap of the group of authors that are the most active in the two data sources. The results indicate that Chinese scholars do not have homogeneous publication patterns: some very productive scholars mostly publish in international (WoS) journals while others prefer to diffuse their research results in national Chinese journals. Disciplines that are most international in scope exhibit a much higher level of overlap than those of the social sciences and humanities. These results suggest that the WoS does not accurately represent Chinese research activities, especially in social science and humanities, but that it also has a relative overlap with the Chinese national scientific literature in the natural and medical sciences.

INTRODUCTION

Over the last 20 years, China's contribution to the world's scientific activity—as measured by its number of Web of Science (WoS) publications—has increased at an impressive rate (Zhou, 2013). While part of this trend might be due to an increase in the number research papers written in English by Chinese researchers (Montgomery, 2013), some Chinese scholars might still prefer to publish their manuscripts in Chinese academic journals (Jin, Zhang, Chen, & Zhu, 2002; Moed, 2002b) only indexed by national Chinese bibliometric databases. Hence, measuring China's research output remains a challenge, as no bibliometric database covers both Chinese and English scientific literature.

Many scholars have, thus, concluded that the WoS is not an appropriate tool to measure Chinese research performance (Guan & He, 2005; Jin & Rousseau, 2004; Zhou & Leydesdorff, 2007), as significant differences were found in the coverage of national and international bibliometric databases (Hennemann, Wang, & Liefner, 2011; Meho & Yang, 2007). While previous work has attempted to explain differences between WoS and Chinese bibliometric databases by looking at journal hierarchies and citation relations (Zhou & Leydesdorff, 2007), or regional publications (Liang, 2003), no research has yet analysed the discrepancies at the level of authors. For instance, little is not known on the extent to which scholars from Chinese institutions publish their articles in international journals, or whether “top” Chinese authors give up publishing papers in Chinese in order to be more visible

¹ This work was supported by the Andrew W. Mellon Foundation.

internationally. A better understanding of those trends might help to explain the differences between the international and Chinese bibliometric databases.

RESEARCH QUESTIONS

The purpose of this study is to 1) compare the WoS with a national Chinese bibliometric database at the level of individual authors, 2) measure the extent of the overlap of the group of authors that are the most active in the two data sources, and 3) compare the results across certain disciplines. Specifically, this study will answer following research questions:

1. In a given discipline, are the top Chinese scholars in terms of the number of publication in WoS also top scholars in the selected Chinese bibliometric database?
2. In a given discipline, are the rankings of Chinese scholars in WoS equivalent to their rankings in the selected Chinese bibliometric database?
3. How does the discipline affect the overlap between top Chinese scholars in international versus national Chinese databases?

In addition to providing a better understanding of Chinese researchers' publication patterns, these results will inform the extent to which the WoS can accurately represent Chinese scientific activity, and how this varies across disciplines.

METHODOLOGY

The Web of Science (WoS) and China Citation Indicators Database (CSI) are used as data sources because of their coverage and representation. WoS is the only bibliometric database covering a century of citation-based indicators for all disciplines, as well as, since 1973, all authors and their institutional affiliations (Moed, 2005). Along these lines, most previous bibliometric studies on China are based on WoS (Zhou & Leydesdorff, 2007). Although there are five major bibliometric databases in China, CSI has the largest coverage and offers author rankings in terms of publications and citations that are not provided by other databases (Zhao, Lei, Ma, & Qiu, 2008).

Inconsistencies are found between WoS and CSI in terms of the discipline classification. WoS assigns journals to 232 subject categories while the CSI classifies journals into 35 fields and 457 subfields. Based on the descriptions of each subject category, 114 one-to-one matches between WoS and CSI were identified. This study analyses data for these 114 disciplines in both WoS and CSI. For each of these disciplines, the number of papers per author was compiled in order to produce ranked lists of top Chinese authors in WoS and CSI. Since CSI only offers bibliometric data from 2000 onwards, the top 100 authors between 2000 and 2014 in the 114 identified disciplines will be retrieved, and form 114 pairs of author groups for comparison. A total of 22,800 records (114 disciplines × 100 authors × 2 databases) were retrieved. The amount of overlap between these 114 pairs is of particular interest since it indicates if, for any of the 114 disciplines, the elite of scholars (i.e. top 100 authors) found in the WoS is the same as the one found in the CSI.

Author name ambiguity is a known threat to the reliability and validity of bibliometric analysis (Moed, 2002a; Strotmann & Zhao, 2012). This is especially the case for Chinese author, as traditional author identification using surname and given name initial(s) cannot distinguish Chinese authors having the same translated English name (Qiu, 2008; Strotmann & Zhao, 2012). In this study, in order to identify the Chinese authors, a Chinese author name database (CAND) has been developed based on the 11,400 records retrieved from CSI. The author's Chinese name, as well as all possible translated English names, and the author's

affiliation were recorded. Chinese authors' English names in WoS were then distinguished by matching the data in CAND. These records were validated manually to produce a list of author name pairs groups in each of the disciplines. Groups of top author groups were then compared in terms of their overlap (i.e., percent of shared authors) and the rankings of individual authors. While the overlap between top authors answers the first research question, the difference in rankings between the groups provides insights on the second research questions. The results of the comparison in different disciplines will answer the third research question.

PRELIMINARY RESULTS

Table 1 lists the 10 specialties selected for the pilot test²; they represent the disciplines of engineering (2), natural sciences, (3) medical sciences (2), social sciences (2) and humanities (1). For each discipline, the lists of top 100 authors in terms of number of publications were retrieved from both WoS and CSI.

Table 1. Number of Chinese papers per discipline, WoS and CSI

Category	Discipline	Papers indexed in WoS (2014)				Papers in CSI (2014)
		From China	Total	China's Ratio	China's Rank	
Engineering	Automation & Control System	2,573	8,261	31.1%	1	20,777
	Engineering, Environmental	2,820	11,240	25.1%	1	21,922
Natural Sciences	Chemistry, Physics	14,730	56,991	25.8%	1	4,133
	Mathematics	4,890	25,296	19.3%	2	12,797
	Optics	8,371	27,125	30.9%	1	2,623
Medical Sciences	Gastroenterology & Hepatology	1,648	11,304	14.6%	2	10,612
	Oncology	8,907	34,960	25.5%	2	42,054
Social Sciences	Management	825	8,528	9.7%	3	3,243
	Education & Education Research	354	9,820	3.6%	9	71,531
Humanities	Linguistics	248	4,795	5.2%	7	3,092

The preliminary results show overlaps between the top 100 authors in the WoS and CSI groups and the Kendall's Tau correlation coefficient between the author rankings. The overlaps between top 10, 20, 50 authors between WoS and CSI groups were also measured as reference.

As Table 2 shown, the overlaps vary from 1% (in *Education & Education Research*) to 21% (in *Engineering, Environmental*). Discipline has a strong effect on the size of the overlap between the groups: while overlaps in engineering and medical science are around 20%, those in the social sciences and humanities are less than 10%. As one might expect, the composition and rank of Chinese scientific elites in WoS and CSI are more closely related in the medical sciences as compared with other categories. The negative correlations were found in three disciplines including two natural science disciplines: *Chemistry, Physics* and *Optics*.

² The CSI data was retrieved for the time period from 2000 to 2015 due to the limited access to the database.

Table 2 The Comparison of 10 Disciplines between WoS and CSI

Category	Discipline	Kendall's Tau Correlation	Overlap in			
			Top 10	Top 20	Top 50	Top 100
Engineering	Automation & Control System	0.111	1	4	8	19
	Engineering, Environmental	-0.134	1	2	12	21
Natural Sciences	Chemistry, Physics	-0.340	0	0	5	18
	Mathematics	0.200	0	1	4	6
	Optics	-0.448	0	0	3	12
Medical Sciences	Gastroenterology & Hepatology	0.162	1	3	10	20
	Oncology	0.517	2	3	8	18
Social Sciences	Management	0.287	0	2	4	8
	Education & Education Research	N/A	0	0	1	1
Humanities	Linguistics	0.300	0	0	0	5

CONCLUSION

This research in progress paper has shown that Chinese scholars do not have homogeneous publication patterns. While some very productive—or *elite scholars*—mostly publish in international (WoS) journals; others prefer to diffuse their research results in national Chinese journals. Unsurprisingly, disciplines that are most international in scope—such as those of the natural and medical sciences—exhibit a much higher level of overlap than those of the social sciences and humanities. Along these lines, we find a higher overlap of top authors between WoS and CSI in disciplines where the proportion of China's contribution to WoS is higher. On the whole, these results suggest that the WoS does not accurately represent Chinese research activities, especially in social science and humanities, but that it also has a relative overlap with the Chinese national scientific literature in the natural and medical sciences. These preliminary results are solely based on the ranking of top authors according to their number of publications, and do not consider the impact of research (i.e. citation). Future work will investigate the Chinese scientific elite according to its average number of citations per paper and total number of citations.

REFERENCES

- Guan, J., & He, Y. (2005). Comparison and evaluation of domestic and international outputs in Information Science & Technology research of China. *Scientometrics*, 65(2), 215-244.
- Hennemann, S., Wang, T., & Liefner, I. (2011). Measuring regional science networks in China: a comparison of international and domestic bibliographic data sources. *Scientometrics*, 88(2), 535-554.
- Jin, B., & Rousseau, R. (2004). Evaluation of research performance and scientometric indicators in China. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of quantitative science and technology research : the use of publication and patent statistics in studies of S&T systems* (pp. 497-514). Dordrecht; London: Kluwer Academic Publishers.

Jin, B., Zhang, J., Chen, D., & Zhu, X. (2002). Development of the Chinese Scientometric Indicators (CSI). *Scientometrics*, 54(1), 145-154.

Liang, L. (2003). Evaluating China's research performance: how do SCI and Chinese indexes compare? *Interdisciplinary Science Reviews*, 28(1), 38-43.

Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of science versus scopus and google scholar. *ASI Journal of the American Society for Information Science and Technology*, 58(13), 2105-2125.

Moed, H. (2002a). The impact-factors debate: the ISI's uses and limits. *Nature*, 415(6873), 731-732.

Moed, H. (2002b). Measuring China's research performance using the Science Citation Index. *Scientometrics*, 53(3), 281-296.

Moed, H. (2005). *Citation analysis in research evaluation*. Dordrecht, UK: Springer.

Montgomery, S. L. (2013). *Does science need a global language? English and the future of research*. Chicago, IL: The University of Chicago Press.

Qiu, J. (2008). Scientific publishing: Identity crisis. *Nature*, 451(7180), 766-767.

Strotmann, A., & Zhao, D. (2012). Author Name Disambiguation: What Difference Does it Make in Author-Based Citation Analysis? *Journal of the American Society for Information Science and Technology*, 63(9), 1820-1833.

Zhao, R.-Y., Lei, j., Ma, R., & Qiu, J. (2008). A comparison among five bibliometric database in terms of citation index. *Information Studies: Theory and Application*, 2008(4), 589-605.

Zhou, P. (2013). The growth momentum of China in producing international scientific publications seems to have slowed down. *Information Processing & Management*, 49(5), 1049-1051.

Zhou, P., & Leydesdorff, L. (2007). A comparison between the China Scientific and Technical Papers and Citations Database and the Science Citation Index in terms of journal hierarchies and interjournal citation relations. *Journal of the American Society for Information Science and Technology*, 58(2), 223.

CHAPTER 10

Gender



Gender in science: a periphery?

Inge van der Weijden*, Emanuela Reale** and Ingeborg Meijer*

i.c.m.van.der.weijden@cwts.leidenuniv.nl; i.meijer@cwts.leidenuniv.nl

Centre for Science and Technology Studies (CWTS), Leiden University, PO Box 905, 2300AX Leiden (The Netherlands)

** emanuela.reale@ircres.cnr.it

Research Institute for Sustainable Economic Growth (IRCRES-CNR), National Research Council, Via dei Taurini 19, Rome, 00185 (Italy)

Words: 1987, excluding reference

ABSTRACT

The academic world has been dominated by men for a long time. However, in recent decades, there have been strides towards gender balance within the pool of higher education graduates. Whilst women were once under-represented at doctoral level, they now made up 47 % of PhD graduates in the EU (European Commission 2015). Although female researchers are improving their position, the process especially in higher positions is rather slow. Is the weak position due to women having in average fewer ambitions in pursuing an academic career? Are career decisions characterized by gendered social closure, structurally disadvantaging women? Or are women weakly represented in high ranks because their male colleagues outperform them? Providing robust answers to the mentioned questions is not simple and even framing sound indicators useful to approximate the phenomenon is still a challenge for scholars dealing with science and technology policy. Therefore, in the context of the conference theme, the question is whether we still could consider gender as a periphery in science. And in particular, what aspects of gender differences in science are in particular affected? Career paths, mobility, performance, leadership and / or funding? These are the topics we want to discuss in a special session of 90 minutes at the STI2016 conference.

The panel session will be organized around a number of (6) invited presentations. There will be 5-8 minutes for a short presentation. Each presentation will be subsequently discussed in a fishbowl setting. This means that we will invite 6 discussants to debate with the presenter on the lessons learned and ways forward for change. For each debate, 3 men and 3 women will be appointed from the audience. Each discussion is required to result in 1 recommendation. We need at least 36 participants to the session in order to have fresh debaters in each round. At the end of the session we will collect six recommendations that can be used in the closing session of STI2016. Further communication of the results may include a special issue for publications and a policy brief to inform European policy and Gender oriented projects.

TOPIC AND PURPOSE

Men have dominated the academic world for a long time. However, in recent decades, there have been strides towards gender balance within the pool of higher education graduates. Whilst women were once under-represented at doctoral level, they now made up 47 % of PhD graduates in the EU (European Commission 2015). Although female researchers are improving their position, the process especially in higher positions is rather slow. Is the weak

position due to women having in average fewer ambitions in pursuing an academic career? Are career decisions characterized by gendered social closure, structurally disadvantaging women? Or are women weakly represented in high ranks because their male colleagues outperform them? Providing robust answers to the mentioned questions is not simple and even framing sound indicators useful to approximate the phenomenon is still a challenge for scholars dealing with science and technology policy. Therefore, in the context of the conference theme, the question is whether we still could consider gender as a periphery in science. And in particular, what aspects of gender differences in science are in particular affected? Career paths, mobility, performance, leadership, funding and even wellbeing? These are the topics we want to discuss in a special session of 90 minutes at the STI2016 conference.

The panel session will be organized around a number of (6) invited presentations. There will be 5-8 minutes for a short presentation. Each presentation will be subsequently discussed in a fishbowl setting. This means that we will invite 6 discussants to debate with the presenter on the lessons learned and ways forward for change. For each debate, 3 men and 3 women will be appointed from the audience. Each discussion is required to result in 1 recommendation. We need at least 36 participants to the session in order to have fresh debaters in each round. At the end of the session we will collect six recommendations that can be used in the closing session of STI2016. Further communication of the results may include a special issue for publications and a policy brief to inform European policy and Gender oriented projects.

PROPOSED ACTIVITIES – 6 PAPER PRESENTATION

1. Emanuele Reale, Antonio Zinilli. IRCRES CNR, Rome, Italy.
emanuela.reale@ircres.cnr.it; antonio.zinilli@ircres.cnr.it

Title: Gender equality and evaluation: do fields of science matter?

The paper assumes that gender bias in R&D activities is a heterogeneous phenomenon, which is influenced by the ideas, values and organizational hierarchies of the different fields of science and the related epistemic communities. Thus, overcoming gender biases needs targeted efforts at national and institutional level that take differences between fields under consideration. To test this assumption, the paper uses indicators derived from the ex-ante assessment of research proposals submitted to competitive funding stream. The empirical base are the PRIN-Projects of National Interest- a funding scheme for curiosity-driven research projects implemented in Italy. Data refer to twelve years (from 2000 to 2011) and four fields of science (Physics, Chemistry, Economics and Social sciences), and combine secondary data of MIUR, Italian Ministry of University and Research, and a survey on 984 university scholars involved as coordinators of the projects and principal investigators of the local research units. The results show relevant field-specific differences in the evaluation practices that affect gender equality, cross-cutting the traditional division between natural sciences and social sciences. The conclusions would infer possible explanations of the existence of vicious relationships between evaluation practices and gender equality in different disciplinary fields, which could impact also research productivity and leadership.

2. Rainer Frietsch, Susanne Bühner, Patricia Helmich – Fraunhofer Institute for Systems and Innovation Research (ISI), Karlsruhe, Germany.
rainer.frietsch@isi.fraunhofer.de ; susanne.buehrer@isi.fraunhofer.de

Title: Scientific and technological output of women and men

In recent studies, we find hardly any difference in the quality (measured by citations) of the output of women and men, both in science (based on bibliometric analyses) and technology (based on patent data). While there still is a gender gap in terms of absolute output and also in terms of certain fields/topics, the quality of the output is not different at all. In addition, we see that women still tend to organize their work differently and seem to be more collaborative than men. However, while our findings support the resource-based line of argumentation, empirical literature exists that comes to contrary conclusions, namely that there is a difference in productivity and also in quality between women and men.

We provide empirical evidence from SCOPUS and PATSTAT on the quality of the output of women and men and its change over past decade. In addition, we argue based on our quantitative analyses that women have a different mode working, for example focusing on different scientific/technological areas and also preferring different team structures. In a multivariate approach, we test differences between men and women controlling for field/topic preferences, country differences, type of affiliation as well as changes over time.

3. Carolina Canibano, Carolina Cañibano, Mary Frank Fox and F. Javier Otamendi
INGENIO (CSIC-Universidad Politécnica de Valencia), Ciudad Politécnica de la Innovación, Camino de Vera s/n, 46022 Valencia, Spain. School of Public Policy, Georgia Institute of Technology, Atlanta, Georgia 30332-0345, USA. Applied Economics I, Universidad Rey Juan Carlos, Paseo Artilleros s/n, 28032 Madrid, Spain.
ccanibano@ingenio.upv.es

Title: Gender and International Mobility of European Researchers

The rise in the cross-country mobility of researchers is explicitly linked to the successful construction of the European Research Area. In addition, mobility is becoming a requisite for promotion in academic careers. This paper focuses on patterns of gender and international mobility using data from the sample of 10547 European researchers who responded to the MORE2¹ survey in 2012. In this sample, 6571 researchers report that they have experienced some type of international mobility of either long-term or short-term. Women represent 36 % of mobile researchers and 48 % of immobile researchers. The survey offers the unique opportunity to study, for a large and international sample of researchers, the potential association between their personal status, among other factors, and their mobility experience. An initial exploratory analysis of the data shows that the likelihood of being internationally mobile is negatively correlated with being a female researcher at all career stages² (R1 to R4) and with having children at all post-doctoral career stages (R2 to R4). In contrast, being in a couple is positively associated with international mobility. The paper also addresses whether

¹ Support for continued data collection and analysis concerning mobility patterns and career paths of researchers. Information and data from the MORE2 project may be found online at <http://www.more-2.eu/>

² The European Framework for Research Careers (European Commission, 2011) structures careers according to four stages: R1, First Stage Researcher; R2, Recognised Researcher; R3, Established Researcher and R4, Leading Researcher.

these results hold across institutional settings, defined in terms of the European Science Foundation classification of EU countries according to differences in prevailing attitudes toward gender roles (EFS, 2013).

4. Carter Bloch, Evanthia K. Schmidt – Danish Centre for Studies in Research and Research Policy (CFA), Aarhus University, Denmark.

carter.bloch@ps.au.dk; eks@ps.au.dk

Title: Gender differences and the role of research grants

Summary: In an increasingly competitive research-funding environment, research grants play an important role for researchers' career advancement. Given this, research grants can constitute a valuable tool in narrowing gender gaps in research. Or, conversely, uneven allocation of funds can exacerbate existing inequalities. Studies reveal that female researchers are less likely to apply for funding and they apply for smaller amounts of resources.

This paper aims at mapping the Danish research funding landscape from a gender perspective and analyzing the gender dynamics in application behavior and among grant recipients. The paper seeks to examine the role of research grants for gender differences by considering the case of the Danish Council for Independent Research. We first examine the Council's focus on gender issues and efforts to narrow gender biases, and developments in success rates for male and female applicants. Thereafter, we will compare outcomes for male and female grant recipients and rejected applicants by scientific area, with particular focus on research performance of grants and subsequent performance following the completion of grant research projects. As part of the analysis, we will employ a matching analysis that ensures comparability of applicants according to a variety of factors, including research performance prior to grant receipt or application.

5. Katia Levecque & Frederik Anseel – Ghent University, Department of Personnel Management, Work & Organizational Psychology, Centre of R&D Monitoring, Ghent, Belgium.

Katia.Levecque@UGent.be; Frederik.Anseel@UGent.be

Title: Gender structured universities and their impact on mental health: a focus on PhD students in Flanders.

A recent study on PhD students in Flanders (Belgium) reported alarming findings: 32% is at risk of a psychiatric disorder. In the current study we scrutinize the mental health impact of the gendered structure of Flemish universities. As in many other regions, they are characterized by increasing numbers of female PhD students, and a consistent pattern of male overrepresentation at both postdoc and faculty level. As a consequence, female PhD students have a higher chance of getting into a different-sex supervisor-student relationship and into a team in which they belong to the minority group. In the current study we assess whether gender of the PhD student, gender composition of the supervisor-student dyad and gender composition of the research team affect PhD students' mental health. Logistic regressions based on a representative survey of PhD students (N=3,659) show that female PhD students have worse mental health, although the gender ratio is lower than expected. Analyses of the supervisor-student dyad suggest that both male and female PhD students do well in both

same-sex and different-sex relationships. For both female and male PhD students, mental health was best in gender-balanced research teams. Belonging to a gender minority group as such did not have health effects.

6. Inge van der Weijden & Ingeborg Meijer, Centre for Science and Technology Studies (CWTS), Leiden University, The Netherlands.

i.c.m.van.der.weijden@cwts.leidenuniv.nl; i.meijer@cwts.leidenuniv.nl

Title: Gender differences in careers after receiving a personal grant

In this paper we study gender differences in careers of scholars who obtained as a postdoc a prestigious personal grant (VENI) from the Dutch research council. We conducted semi-structured interviews. 4 clusters, or typologies of careers could be discerned: 1) Dedicated tenured researchers with teaching tasks; 2) Dedicated tenured researchers with clinical tasks; 3) Part-time tenured researchers with major clinical tasks and/or coordination & management tasks, or even without any remaining research tasks in an academic environment; 4) Postdoc hoppers. While at the moment of granting the gender balance was even, the distribution over the clusters six years later is not. The classic research profile is dominated by men (65%), and in the mixed and post doc profile women are dominant (80%). This suggests that policies aiming for excellence, like the VENI, do not change the current gendered system and discourse of meritocracy in academia. Interestingly, in the classic cluster, the women do benefit from specific programmes for women in science, advancing their careers towards associate professor relatively faster. However, in the mixed cluster, women explained that they reduced their research effort because they aimed for a more diverse job that tapped into a wider set of skills and capacities, stating that research turned into a monoculture of publications and grant applications.

INTENDED AUDIENCE

The expected audience to the special session is at least half of the conference attendants, assuming that all women should know and debate about this. Alternatively, the other half of the conference participants, the male researchers, should be attending the session in order to further increase awareness about explicit and implicit bias. Otherwise, we would be happy to accommodate 40 attendees.

The organizers have been working in the field of STS and STI since several years. Last STI conference (2015 in Lugano), we agreed on the fact that we missed special attention for the gender issue. With the periphery focus, we thought we could make a nice contribution.

NOVELTY

Gender equality in research and innovation is one of the core dimensions of Responsible research and innovation. It requires permanent attention and hence a special session on this topic is timely and necessary.

REFERENCES

Commission E (2015) *She Figures 2015: Gender in Research and Innovation* (Publications Office of the European Union, Luxembourg).



What drives the gender gap in STEM? The SAGA Science, Technology and Innovation Gender Objectives List (STI GOL) as a new approach to linking indicators to STI policies

Ernesto Fernández Polcuch*, Martin Schaaper**, Alessandro Bello***

*e.fernandez-polcuch@unesco.org

Chief of Section, Science Policy and Partnerships, UNESCO, 7 Place de Fontenoy, Paris, 75007 (France)

**m.schaaper@unesco.org

Programme Specialist, Science, Technology and Innovation Unit, UNESCO Institute for Statistics (UIS), 5255 avenue Decelles, Montreal (Canada)

***a.bello@unesco.org

Project Officer, SAGA, UNESCO, Luis Piera 1992, Montevideo, 11000 (Uruguay)

ABSTRACT

There is a large imbalance in the participation of women in Science, Technology, Engineering and Mathematics (STEM) fields across all of Latin American countries despite the fact that the region has one of the highest proportions of female researchers worldwide (44% according to UIS statistics). Female researchers face persisting institutional and cultural barriers, which limit the development of their careers and constrains their access to decision-making positions. In this framework, UNESCO has launched the STEM and Gender Advancement (SAGA) project, which has for objective to address the gender gap in STEM fields in all countries at all levels of education and research as well as to promote women's participation in science.

SAGA is a global UNESCO project with the support of the Swedish Government through the Swedish International Development Cooperation Agency (Sida).

One of the outcomes of this project is the SAGA Science, Technology and Innovation Gender Objectives List (STI GOL), which is an innovative tool that aids in the identification of gaps in the policy mix. Additionally, the STI GOL configures the conceptual backbone of the SAGA project, by linking gender equality STI policy instruments with indicators.

By using the STI GOL, and identifying the gender gaps, policy-makers will be able to implement evidence-based policies in STEM fields.

The SAGA STI GOL is a new and innovative way of contributing to the development of effective gender sensitive policies in STI fields, both in education and in the workplace. Likewise, it enables the categorization of STI policies and instruments, with the objective of identifying gaps in the policy mix and aid in the creation and design of evidence-based public policies to promote gender equality.

Keywords: STEM, Gender policies, STI Policies, Women in Science

Despite the remarkable gains that women have made in education and the workforce over the past decades, progress has been uneven. According to estimates by the UNESCO Institute for Statistics (UIS), only 28% of the world's researchers are women. While a growing number of women are enrolling in university, many drop out at the highest levels required for a research

career. Furthermore, women are still underrepresented in the fields of science, technology, engineering and mathematics (STEM), both in the number of graduates (especially at the Ph.D. level), and in the research profession (see for example [UNESCO Science Report: Towards 2030](#) or the UIS [Women in Science](#) visualisation), with the gender gap particularly apparent in disciplines such as mathematics, engineering and computer science.

Although the development of STEM fields is widely regarded as beneficial to the expansion of national economies, the underrepresentation of women in STEM represents the loss of a critical mass of talent and ideas. Therefore, reaching gender equality in STEM implies encouraging further the participation of women and girls at all levels of education, and providing equal opportunities for scientists and engineers throughout their careers. Achieving gender equality is an overarching UNESCO priority, both as a matter of human rights and in order to enhance countries' STI capacities. It is also crucial to achieve the 2030 Agenda for Sustainable Development, in particular, Sustainable Development Goal 5: "Achieve gender equality and empower all women and girls", but also all other SDGs relying on STI capacities.

The current gender imbalance in STEM is partly a consequence of long-term implicit and explicit policies and policy instruments put in place at various levels, inside and outside the STEM system (government, funding agencies, higher education institutions, research centres, inter alia), in addition to social and cultural factors. However, no guidelines exist at the global level to assist governments and policy-makers in the creation of policies aimed at ensuring the participation of young girls and women in STEM.

Moreover, the lack of data from which to draw useful indicators and proceed to analytical studies can obstruct the design, monitoring and evaluation of STI policies aimed at gender equality. Indeed, effective STI policies need to be evidence-based to allow for the design of policies adequately addressing the gender imbalance in STEM. Considering this, there is an urgent need to define and collect better policy-relevant indicators on all aspects of women in STEM to address current and potentially arising issues in the future.

To bridge these gaps in information and tools, a new methodology has been elaborated to support policy-makers worldwide in setting up, implementing, monitoring and evaluating gender equality policies in STI: the SAGA Science, Technology and Innovation Gender Objectives List (STI GOL).

The STI GOL is an outcome of SAGA (STEM and Gender Advancement), a global UNESCO project supported by the Swedish International Development Cooperation Agency (Sida), which aims at identifying gaps and best practices for designing and improving public policies promoting gender equality in STEM. Its mandate is to develop new and better indicators to provide tools for evidence-based policy-making in STI thereby addressing existing and emerging issues arising during evidence-based policy-making due to the lack of adequate information.

The STI GOL approach consists of relating indicators and policies to gender objectives and identifying gaps in the policy-mix. It offers a methodology to assess policies and is based on seven objectives, which carry important policy impacts:

1. Social norms and stereotypes
2. Primary and secondary education
3. Higher education
4. Career progression
5. Research content and practice
6. Policy-making processes
7. Entrepreneurship and innovation

These seven areas configure the first level of the STI GOL. A second level provides breakdowns that allow for further classifying policy instruments and indicators.

The STI GOL is the product of an analytical and conceptual clustering of STI gender-related policies and instruments. Subsequently, the list of objectives was presented and reviewed by a team of prominent international experts in STI policy and in gender equality, from organizations with a stake in the subject. It does not necessarily constitute a formal “classification”, since policies and instruments may be placed under more than one STI Gender Objective. Each objective seeks to answer the question, what drives the gender gap in STEM?

Many factors contribute to the gender gap in STEM, including access to equal opportunities in education and career development and socio-cultural barriers. Therefore a holistic and multi-pronged approach to close the gaps is needed. The seven objectives of the STI GOL, intrinsically interrelated, aim at encompassing all aspects of gender equality in STI policy making. It is of paramount importance to change perceptions, attitudes, behaviour, social norms and stereotypes towards women in STEM in all societies as well as engage girls and young women in STEM in primary and secondary education. Women also face challenges along the STEM career, thus it is essential to promote the attraction, access to and retention of women in STEM higher education at all levels while likewise promote gender equality in career progression in STEM.

Addressing challenges associated with retaining women in STEM also means promoting the gender dimension in research content, practice and agendas by conducting gender sensitive analyses in the research process, when developing concepts and theories, collecting and analysing data and using the analytical tools that are specific to each scientific area.

Crucial is also ensuring the gender balance in STEM-related policy design, through the promotion of “positive action” to expand women’s participation in policy and decision-making processes in all fields of society, as well as ensuring gender mainstreaming and prioritization of gender equality in the design, monitoring and evaluation of STEM related policy. Additionally, it is essential to promote gender equality in science and technology-based entrepreneurship and innovation activities.

The SAGA STI GOL allows for a wide-ranging map and classification of STI policy instruments, while also assessing the availability of indicators in each area and identifying objectives for which indicators do not yet exist. This demonstrates the extent of available statistical information and aid in overcoming the lack of reliable information. Therefore, the STI-GOL will improve the design, monitoring and evaluation of public policies to promote

gender equality in STEM. Furthermore, this mapping will also highlight gaps in the policy mix, driving STI policy- and other decision-makers towards the development and implementation of new policies and instruments at the country level, particularly when the specific STI gender objective is shown to be in need of intervention by the corresponding indicators. This link between data and solid analysis is critical for policymakers as they look for ways to integrate these advances into policy design.

The SAGA STI GOL looks more closely at how feasible it is to make a direct link between STI gender-related policies and indicators. By linking them, it becomes easier to monitor and evaluate the impacts policies have and provides evidence for improving gender equality in STEM. An initial matrix attempt is currently being undertaken and it will be present in various pilot countries around the world.

A number of different concepts are used in the STI GOL, including most notably “STEM”, “S&E”, and “STI”:

- STEM – Science, Technology, Engineering and Mathematics – is used to characterize the corresponding fields of knowledge and study.
- S&E – Scientists and Engineers is used when dealing with professions, most frequently carried out by graduates of STEM Higher Education careers.
- STI – Science, Technology and Innovation – is used when referring to policies.

The new methodology will enable countries to improve gender-related-policies by highlighting areas in their policies that do not support the promotion of women’s participation in STEM. Using the STI GOL to create evidence-based policies and programmes, countries will be in a better position to eliminate obstacles and increase young girls and women’s participation in STEM.

The SAGA STI GOL was implemented for the first time in the case of Latin America, which was achieved by following two methodologies. First, an inventory of instruments identified in the analysed databases was carried out. This allowed to make a first analysis of gaps in the policy mix. Then, a research on ministries and government institutions, which cover each of the seven areas or gender objectives, was made in order to identify other instruments which had not been added to the different databases. By doing so it was possible to carry out a more complete and up-to-date analysis identifying the gaps in the different gender objectives.

In conclusion, the SAGA STI GOL is an innovative tool for implementing effective policies in STI, both in education and in the workplace. In fact, the STI GOL will be used in upcoming work on gender in STEM policies and indicators by UNESCO. The Inter-American Development Bank has already committed to applying it in its project “Gender Gaps in Science, Technology and Innovation in Latin America and the Caribbean”. The STI GOL is continuously evolving towards building becoming an efficient tool for setting up, implementing and monitoring gender policies.

The STI conference is the appropriate platform to present this new methodology because of the type of forum and audience attending the event. The exchange and the dissemination of best practices and innovative methodologies to address gender imbalance, which are evident around the world, are important to improve mainstreaming gender considerations into research, innovation, and policy.



Picking the best publications to showcase graduate courses: Do institutional mechanisms reinforce gender differences?¹

Jacqueline Leta* and Guillaume Cabanac**

* jleta@bioqmed.ufrj.br

Instituto de Bioquímica Médica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 20941-900 (Brazil)

** guillaume.cabanac@univ-tlse3.fr

Computer Science Department, IRIT UMR 5505 CNRS, Université Toulouse 3, 118 route de Narbonne,
F-31062 Toulouse (France)

ABSTRACT

Previous studies performed by our research group have brought to light the concept “scientific capital” developed by Bourdieu to characterize the vertical segregation framework in Brazilian science, specifically Brazilian graduate programs, which is the main piece of the country’s S&T system. The present study still focuses on gender differences in Brazilian graduate programs but it turns attention to their institutional assignments. Among all information sent annually to Capes evaluation process, heads of graduate programs have to send the top five publications of the year. Considering the institutional relevance of this set of publications, the present study aims to identify whether an institutional mechanism, as the choice of the best publications of the graduate program by the heads, promotes gender equality or reinforces discrepancies in Brazilian academia. Preliminary results, performed upon official data of teacher-researchers performance affiliated to Federal University of Rio de Janeiro, suggest that males rather than females more are more represented in this selective set of publications.

INTRODUCTION

Some decades ago, we have witnessed the flourishing of a new field of knowledge devoted to gender studies in science activities, generally called “women in/and science” or “gender in/and science”. In recent years, this field has displayed both an impressive growth in the number of scientific publications and a diversity of areas involved on it (Dehdarirad, Villarroya & Barrios, 2015).

In the social science literature on women and science issue, the identification of social and institutional factors involved in the success of men and women in scientific careers appears as one of its main targets (Schienbinger, 2001). In line with this approach are the studies on author productivity that, despite the lack of a consensus, mostly reveal an uneven scene between men and women, where men usually present higher rates of papers and citations (e.g., see Long, 1992; Prpić, 2002; Larivière et al., 2013). Such gender differences in productivity may represent a disadvantage for women and, consequently, for their career advancement (e.g., see Long, Allison & McGinnis, 1993; van Arensbergen, Weijden & Besselaar, 2012).

Previous studies developed by Leta’s research group (Leta et al., 2013; Olinto & Leta, 2015) have brought to light the concept “scientific capital” developed by Bourdieu (1997) to

¹ This work was supported by CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Franco-Brazilian Research Group on Web Science (CNRS GDRI WebScience).

characterize the vertical segregation framework in Brazilian science, specifically the graduate programs, which is the main piece of the country's Science and Technology (S&T) system. The goal was to map whether different academic tasks were evenly distributed between male and female teacher-researchers, that is, those affiliated to a Brazilian graduate course. Considering the complexity of graduate programs environment, the expectation was to find Brazilian male teacher-researchers showing higher burdens of time consuming in tasks, who promote and reflect higher levels of "scientific capital", such as publishing in top-ranked journals.

It is important to highlight that there is a strong inter-relationship between Brazilian graduate programs and S&T activities. Since the beginning of the 1990's, graduate programs are regularly evaluated by Capes, an agency of the Ministry of Education. Different indicators are considered in this evaluation but, depending on the field of the graduate program, the number of publications (especially with international visibility) is the main criterion for getting the higher grades in the evaluation. In this scenario, Brazilian teacher-researchers have to cope with both their workaday roles (e.g., teaching undergraduate and graduate courses) and with the pressure to publish original research.

The present study still focuses on gender differences in the Brazilian graduate programs but it turns attention to their institutional assignments. Every year, heads of graduate programs are required to organize and send to Capes hundreds of academic and scientific information related to each different tasks performed by teacher-researchers who are under their responsibility. The quality and reliability of the whole set of information sent to Capes may result in a higher grade for graduate programs, which in turn means higher amounts of resource for the program. Among the information sent to Capes are the best five publications of the year, which are expected to be representative of the performance of the whole staff during the year. In most of the cases, the decision to choose the best five publications is the sole responsibility of the heads of graduate programs.

Considering the institutional relevance of this set of publications, the present study tackles the following research question: How are male and female teacher-researchers featured in the best publications of graduate programs? Thus, the study aims to identify whether an institutional mechanism (i.e., the choice of the best publications of the graduate program), promotes gender equality or reinforces discrepancies in Brazilian academia. Preliminary results from this first essay suggest that program heads tend—consciously or not—to indicate more male-authored publications. Nevertheless, such observation needs to be corroborated with some additional analyses.

METHODS

As input of a document analysis technique, the main source of information was a form named "PB - Produção Bibliográfica," which contains the list of all publications published in a given year as well as the indication of the five best publications per program. This form (one out of 11 in total) is an official document elaborated by Capes and it is part of the set on documents that each graduate program is required to submit to Capes for the annual evaluation process. All forms are accessed through the following URL:

<http://conteudoweb.capes.gov.br/conteudoweb/CadernoAvaliacaoServlet>.

For the present study, we downloaded the PB forms available in PDF format for 91 programs in 2009 and 100 programs in 2012 registered by Federal University of Rio de Janeiro (UFRJ,

short name in Portuguese). UFRJ is the oldest and largest public university in Brazil supported by funds from federal government. We then extracted the text of the best five publications appearing under the “Trabalho Completo - Qualis” headings. We focused our attention on the first author of each of these publications addressed as “docente” (Portuguese word for teacher). About 700 teacher-researcher names (i.e., *docentes*) were then identified in the bibliographic entries.

A particular difficulty for this study — as in all studies about women in science — is the availability of information about the scientist’s sex. The PB-form does not mention the author’s sex of best publications. In addition, the PB form identifies authors by their linkage to the graduate program and only those identified as “teacher” were considered. Eventually, each “teacher” was manually tagged with a sex based on the annotator’s knowledge and on information provided online (e.g., Lates CV, personal webpage). We were unable to identify the sex of teachers in 15 best publications only (8 in 2009 and 7 in 2012).

Information about the grade of each graduate program awarded by Capes in 2009 and 2012, as well as about the type of publication was also added (semi-automatically) to the original file.

After data cleaning and duplicate removal, data on 90 and 97 graduate programs and 366 and 384 best publications in 2009 and 2012, respectively, were the basis for this case study, since they refer to a single institution, the UFRJ. It’s noteworthy that data for 2009 and 2012 are available online and refer to the last years of the Capes triennial evaluation processes; 2015 is not yet available.

RESULTS

Among the 735 best publications of UFRJ’s graduate programs (those which had the identification of author’s sex, the teacher), 60.4% are authored by men as first “docente” author ($n = 444$) and 39.6% by women as first “docente” author ($n = 291$). The same distribution is found when the year of the best publications is considered (Table 1).

When checking the total amount of male and female teacher-researchers registered at UFRJ’s graduate programs in 2009, the distribution is as follows: 56.5% men ($n = 1,318$) and 43.5% women ($n = 1,016$). Although it is not the best comparison, the distribution of total amount of male and female among graduate programs’ staff suggests that women are slightly underrepresented among the authorships in the set of publications classified as the “best” of graduate programs.

Table 1: Number and percentage of male and female teacher-researchers as authors in the best publications of UFRJ’s graduate programs, 2009 and 2012.

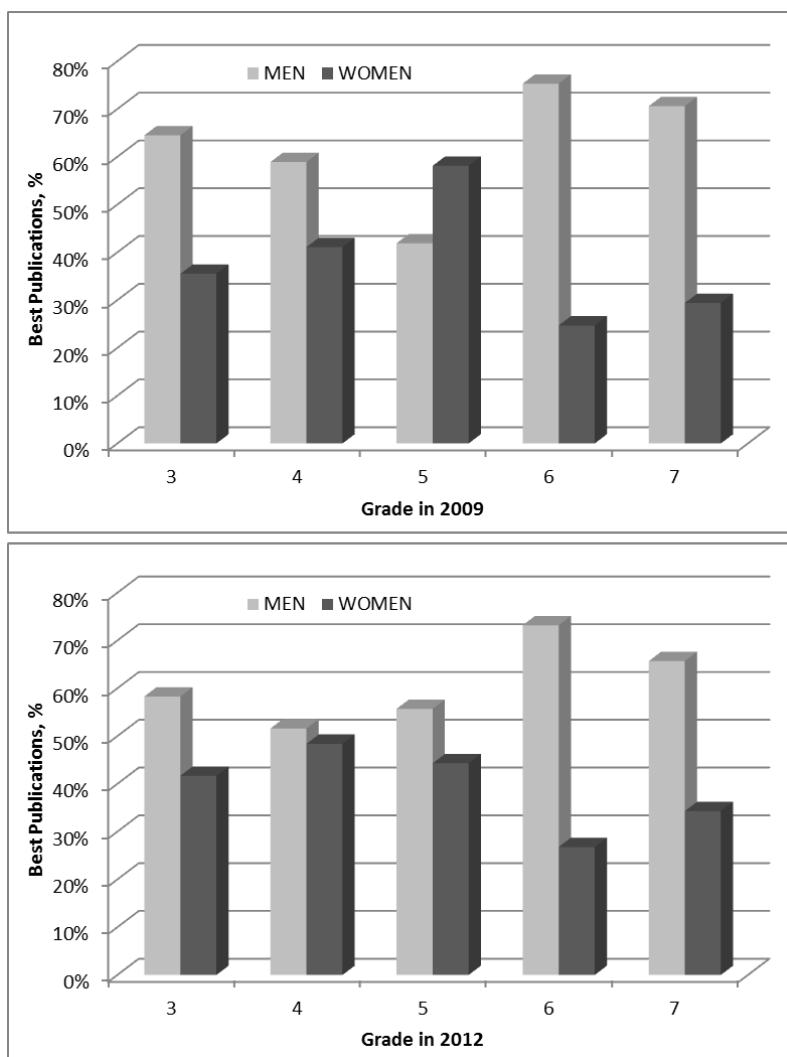
Gender	2009	2012	Total	2009 (%)	2012 (%)
Men	218	226	444	60,9	59,9
Women	140	151	291	39,1	40,1
Total	358	377	735	100,0	100,0

The 90 and 97 UFRJ’s graduate programs registered in Capes in 2009 and 2012, respectively, were evaluated according to their academic and scientific performance in the respective year.

The grades were recorded in a scale from 3 to 7. Usually grade 3 is granted to younger programs. Along the evaluation period, it is expected that all the newest programs reach higher grades as far as grades 6 and 7, which are granted to more established programs with highest performance.

The share of male and female teacher-researchers in the best publications of UFRJ’s graduate programs considered also the grade of the program he/she were linked to, as it is shown in Figure 1. A first insight is that, with the exception of year 2009, grade 5 programs, the chance to have a female-authored publication among the best publication is always lower than a male-authored. This chance increases among lower grade programs but it reduces dramatically among the programs with the highest performance, that is, among grade 6 and 7 programs.

Figure 1: Percentage of male and female teacher-researchers as first “docente” authors in the best publications of UFRJ’s graduate programs by Capes performance grade, 2009 and 2012



Considering only grade 7 programs, women-authored publications (as first “docente” in the byline) represent about 30% of all best publications. The eleven graduate programs included,

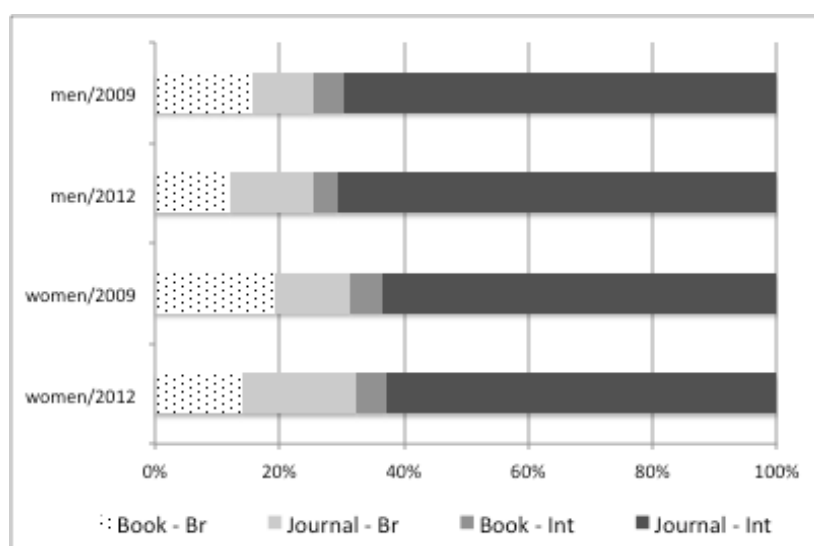
in 2009, in this very selective set of programs featured 459 teacher-researchers, 154 of which were women (33.4%). However, five out of the eleven programs were in engineering, an area where women are ever underrepresented. Together these programs summed 171 teachers, being 21 women (12.3%).

Hence, although Figure 1 points to an underrepresentation of women as first “docente” authors of the best publications, it seems indeed that the share of women in this set of publications is in accordance to the share of women among the whole staff of teacher-researchers linked to UFRJ’s grade 7 programs.

A final aspect investigated in the 735 best publications of UFRJ’s graduate programs was the type of publications. Each publication was classified into one of the four main groups: (1) book or chapter edited in Brazil, (2) book or chapter edited abroad, (3) article published in a Brazilian journal and (4) article published in an international journal. The expectation in performing this analysis was to observe whether the most relevant publications of male and female teacher-researchers have similar or different targets in terms of venues, readership, and visibility.

The distribution of each type of publication among male and female teacher-researchers total publications in 2009 and 2012 is shown in Figure 2. As it can be seen, independently of the sex of the “teachers,” the most relevant publications of UFRJ’s graduate programs are articles published in international journals. In other words, the choice of the best publications prioritises publications geared to peers abroad for both males and females. Such trend may be a result of Capes annual evaluation, which increasingly incentivises Brazilian scientific community to publish in international journals (Leta, 2012).

Figure 2: Distribution (%) of publication type by male and female teacher-researchers as first “docente” authors in the best publications of UFRJ’s graduate programs, 2009 and 2012



Although data apparently suggest a uniform performance in terms of types of publication indicated to male and female teacher-researchers by the heads, a more detailed look in the data indicates a slight tendency for men to increase the share of articles in international

journals. On the other hand, women tend to increase the share of articles published in Brazilian journals.

DISCUSSION

Considering the research question “how are male and female teacher-researchers represented in the best publications of graduate programs?”, the set of preliminary results shown in this paper suggests that males rather than females are increasingly represented as first “docente” authors in this selective set of UFRJ’s publications. Since the choice behind picking the best publications is mostly a decision of the head of graduate program, this institutional decision is, apparently, reinforcing gender discrepancies in our case study, UFRJ’s graduate programs.

Nevertheless, we do believe such an observation needs to be corroborated with some additional analyses, for instance: to compare the ratios of male and female both in the best five publications and in the total corpus of authored papers of the UFRJ graduate programs. Other complementary analysis would be to assign 1/n authorship credit to each author that is, to proceed a fractional counting.

Next steps include the development of these analyses, as well as others to better characterize the top five publications (such as the impact factors of journals). Our aim is to provide a better understanding of the rationale behind how these publications are picked. To the best of our knowledge, this specific dataset has not been studied in women in science studies to date. Hence, although it deals with Brazilian academia only, its originality may bring new insights about institutional mechanisms that push vertical segregation, forcing women to assume mostly the periphery in Brazilian science.

REFERENCES

- van Arensbergen, P., Weijden, I. & Besselaar, P. (2012) Gender differences in scientific productivity: A persisting phenomenon? *Scientometrics*, 93, 3, 857–868.
- Bourdieu, P. *Les usages sociaux de la science : pour une sociologie clinique du champ scientifique*. Paris: INRA, 1997.
- Dehdarirad, T., Villarroya, A., & Barrios, M (2015) Research on women in science and higher education: a bibliometric analysis. *Scientometrics*, 103, 3, 795–812.
- Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature*, 504, 7479, 211–213.
- Leta, J., Olinto, G., Batista, Pablo, D. & Gender & Borges, E.P. (2013). Academic Roles in Graduate Programs: Analyses of Brazilian Government Data. In: 14th International Conference of the International Society for Scientometrics and Informetrics, 2013, Vienna, Austria. *Proceedings of the 14th Conference of ISSI*. Vienna: AIT Austrian Institute of Technology GmbH, 2013. v. 1. p. 796-810.
- Leta, J. (2012) Brazilian growth in the mainstream science: The role of human resources and national journals. *Journal of Scientometric Research*, 1: 44-52.
- Long, J.S. (1992). Measures of sex differences in scientific productivity. *Social Forces*, 71, 1, 159-178.
- Long, J.S., Allison, P.A. & McGinnis, R (1993) Rank Advancement in Academic Careers: Sex Differences and the Effects of Productivity. *American Sociological Review*, 58, 5, 703-722.

- Olinto, G., Leta, J. Scientific Production in Brazilian Research Institutes: Do Institutional Context, Background Characteristics and Academic Tasks Contribute to Gender Differences?. In: ISSI 2015 - International Conference of the International Society of Scientometrics and Informetrics, 2015, Istanbul. *Proceedings of 15th International Society of Scientometrics and Informetrics Conference*. Istanbul: Bogaziçi University Printhouse., 2015. v. 1. p. 673-683.
- Prpić, K. (2002) Gender and productivity differentials in science. *Scientometrics*, 55, 1, 27–58.
- Schienbinger, L. O feminismo mudou a ciência? Bauru: EDUSC, 2001.



What factors influence scientific and technological output: The case of Thailand and Malaysia¹

Catherine Beaudry* and Carl St-Pierre**

* catherine.beaudry@polymtl.ca

** carl.st-pierre@polymtl.ca

Department of mathematics and industrial engineering, Polytechnique Montreal, P.O. Box 6079, Downtown office Montreal, Quebec, H3C 3A7 (Canada)

Keywords: research output, age, gender, funding, collaboration

ABSTRACT

The paper aims to examine the factors that impact scientific outputs and technological outcomes in two Asian countries, Malaysia and Thailand. Using a survey instrument sent to young scientists in these two countries, we find that devoting a higher proportion of time to teaching, which we associate with career maturity, raising a greater proportion of research funds from international sources and collaborating more often at the national level are the main factors that influence research output. In addition, the survey shows that men are slightly more prolific than women in terms of research output, but the difference is not statistically significant. Moreover, once we account for a variety of factors that influence scientific production, our research does not give credence to the common argument that female researchers are less prolific, with one exception, however, women who have more children are less productive than their male counterparts.

INTRODUCTION

Becoming fully established as a member of the academic profession and pursuing access to a permanent position is a critical career goal for many young scientists and researchers all over the world. Their career paths, which is increasingly mobile and international, is also strongly shaped by local and national institutions and highly dependent on scientific production and impact. Several factors influence research performance, which ultimately contributes to building a research career. A number of these factors are socio-demographic, age and gender for instance, others are related to the choices made by the researchers, collaboration and funding spring to mind in this regard.

To get a better understanding of what obstacles and opportunities influence scientific production and thus shape young scientist careers, a questionnaire was sent to young scientists in Thailand and Malaysia. This short paper hence examines the factors identified in the survey

¹ This work was supported by the National Science Technology and Innovation Policy Agency (NSTDA) of Thailand, the Federal Ministry of Science and Education (Bundesministerium für Bildung und Forschung – BMBF) of Germany and the Canada Research Chair on the Creation, development and commercialisation of innovation. We acknowledge the work of Prof. Futao Huang and Dr. Orakanoke Phanraksa, PI and co-PI with Prof Beaudry on the grant. We are grateful for the precursor work of Irene Friesenhahn, the substantial coordination work of Johannes Geffers on the questionnaire, as well as the insights from all the Global Young Academy (GYA) members who commented on the questionnaire. Finally, we are indebted to Laurence Solar-Pelletier for taking a very long paper and reducing it to less than 3,000 words. The opinions expressed do not represent the views of NSTDA, BMBF nor of the GYA. The authors are solely responsible for the contents.

that may be associated with a greater scientific and technological productivity. The next few paragraphs briefly survey the pertinent literature and propose hypotheses that will be further tested using appropriate regression models.

Age

The relationship between the age of researchers and their scientific productivity or scientific impact has been under scrutiny for a great number of years now (for a review of the topic, see Feist, 2006). Lehman (1953) demonstrated that major contributions are likely to occur when scientists are in their late 30s or early 40s, and thereafter decline rapidly. Since this seminal paper, the literature on the relationship between ageing and research productivity appears to be separated into two groups, each reporting opposite findings. Some claim that scientists conduct their best work while young (Einstein, Newton and Gauss are obvious examples), while others argue that knowledge matures with age (Plank, Braun and Cram were in their 40s when they formulated their theories). The first group generally advocates that younger researchers are more productive and more likely to be cited than their older colleagues (Over, 1988) and that extraordinary achievements tend to occur before the age of 40 (Dietrich & Srinivasan, 2007). In contrast, the second group of studies argues that it is not the younger researchers, but the mid-career- and older researchers, who produce the most research and have a greater scientific impact. With this in mind, our first hypothesis proposes that:

H1 (Age): Older young scientists are more productive in terms of research output.

Gender

A vast literature highlights the poor research performance of women in relation to that of men. On average, women publish fewer papers than their male colleagues (Fox, 2005). Some scholars have, however, noted a narrowing of the gap in the publication differences between gender, as the population of female scientists increases (Abramo et al., 2009), and no gender effect on scientific productivity has been found for certain fields. In addition, women seem to be less productive in the first decade of their career, but are more productive afterwards (Long, 1992). A smaller proportion of women benefit from research funds, but both men and women receive grant amounts proportional to the number of submitted proposals at NIH and NSF (Fox, 1991). A number of explanations for these discrepancies have been put forward over the years. For instance, opportunities for women to collaborate are significantly less than those for men when women have young children and are therefore less mobile. Indeed, Larivière et al. (2013: 213) found that “female collaborations are more domestically oriented than are the collaborations of males from the same country”. It would therefore seem that childcare, the age of the children (Fox, 2005) and the lack of research collaboration are the main obstacles to increase productivity (Kyvik & Teigen, 1996). In fact, childcare affects the productivity of women but not that of men (MIT, 1999). Some scholars advance that there are broad gender inequalities regarding access to research funding and equipment. Women often work in universities with a lesser research intensity. Furthermore, women devote more time to teaching and administrative duties than men (DesRoches et al., 2010) and specialise less than men (Leahey, 2006). Our second hypothesis is therefore:

H2 (Gender): Female researchers are less prolific in terms of scientific output.

Funding

Some studies have shown that better-funded scientists are more frequently cited and more productive than less-funded scientists (Beaudry & Allaoui, 2012): the granting of research

money further acts as a signal that attracts additional funding in subsequent years; research financing has a strong positive impact on the number of scientific articles published; and specific grants add one additional publication within the five years subsequent to the attribution of the grant (Jacob and Lefgren, 2007). Furthermore, industrial R&D contracts and funding from private sources have an impact if they represent a small proportion of total funding: “R&D contracts with industry and academic research activities have synergistic effects on scientific production, but only when R&D contracts account for a small percentage of a researcher’s total funding, otherwise, there are decreasing marginal returns to scientific output” (Manjarrés-Henríquez et al., 2009: 799). In this regard, other researchers found a positive effect of philanthropic funding coming from not-for-profit organisations. We propose that:

H3 (Funding): Researchers with a higher proportion of funding from (a) public national organisations will also generate more scientific output, while researchers with a higher proportion of funding from (b) private organisations or (c) philanthropic organisations will generate more technological output.

Collaboration

Networking and collaborating are both beneficial towards scientific production. In addition, collaboration can become a powerful lever to raise funds (Daniel et al., 2003), and consequently, scientific collaboration and research funding are intrinsically intertwined. Multi-project research centres encourage researchers and their universities to collaborate more efficiently, thereby leading to a more efficient use of the available diversity of resources of a physical, human and/or financial nature (Zucker et al., 2007). The fact that most papers are now written in collaboration may contribute to reducing the gender differences. Kyvik and Teigen (1996) identified the lack of research collaboration as one of the main obstacles to increasing research productivity. We would therefore expect collaboration to have a positive impact on research production, but that because women work in smaller or more localised teams their research may be less numerous. Our last hypothesis is therefore:

H4 (Collaboration): Researchers who collaborate will also generate more research output.

DATA AND METHODOLOGY

Data

The study is based on a questionnaire sent to all young researchers in Thailand and Malaysia, with a reminder two weeks later. Team members and colleagues² contacted the main research institutions of these countries, both public and private, in order to gain access to email lists of young researchers in these institutions. The questionnaire was launched in two phases using a convenience sampling technique; 218 responses were collected in April-June 2015, and 534 responses were collected in July-September 2015. This second wave of responses suffered from a significant respondent fatigue problem and thus resulted in only 325 valid responses. As a consequence, tests to compare the two samples were performed, but showed no significant differences between the two groups for the main variables of interest. Table A1 in the appendix describes the variables used in the model and a comparison between genders.

Model

We have identified two potential dependent variables: traditional research output measured by the number of articles, book chapters and conference presentations, as well as the number of

² Acknowledged in the first footnote.

pending and granted patents. These dependent variables being highly skewed, the empirical distribution is better represented by a log normal distribution. While the former follows a normal distribution once we have taken the natural logarithm of the variable, and can thus be analysed using Ordinary Least Squares (OLS), the latter comprises a significant number of zeros and has therefore been estimated using left-censored Tobit regressions. Once all missing values are accounted for, we are left with a sample of 338 observations on which the regression analysis was performed.

REGRESSION RESULTS

Table 1 presents the OLS regression results for the various factors that are associated with a higher number of articles, book chapters and conference presentations, while Table 2 presents the Tobit regression results on the number of patents³. During the course of our study, we compared „real“ age with „PhD“ age and chose the latter as yielding better and more robust results. As a standalone variable, neither variable was ever significant, which is not surprising considering the fact that our sample is composed mainly of young scientists. The effects that other scholars are measuring on vast cohorts or differently aged scientists are simply non-existent with a more homogeneous cohort. Only when interacted with gender and with the proportion of hours dedicated to various tasks (column Art-4), or with foreign collaboration (column Pat-12) was „PhD“ age significant. As researchers age, only a higher proportion of time devoted to research tasks has a positive impact on scientific output, more time dedicated to teaching or to applying for grants has a negative effect.

Similarly, gender as a standalone variable is not significant. Gender, however, moderates the relationship between research output and „PhD“ age, the number of children, the proportion of hours devoted to research, foreign collaboration, and mobility. As such, our research cannot say that female researchers are less prolific once we account for a variety of factors that influence scientific production. For instance, although young women produce fewer publications in their early career, our results suggest that as they grow older, they make up for this low performance (Figure 1 illustrates the results of column Art-10 in this regard). Congruently to what is generally found in the literature, women who have more children (column Art-2) are less productive than their male counterparts. Having children is, however, associated with a degree of maturity that we do not successfully capture with age. Our results show that men with children (column Art-3) are more productive, followed by women with children. Furthermore, female scientists who collaborate with foreign partners do not reduce their technological output, i.e. patents (column Pat-12) as much as men do when they collaborate with foreign colleagues.

Our funding variables only highlight the importance of private funding and of international funding for research output. Contrarily to most studies, we do not have access to the specific amounts of funding raised by individual researchers but only to the proportion of funding from each source. We would therefore not expect to replicate most results from the literature. Private and international funding matter more for classic research output (in Table 1). Surprisingly, private funding has no impact on technological output (in Table 2). For all four categories of output, however, international funding has a strong positive relationship. We therefore suggest that the funding model that brings consensus in the literature may not be

³ Various transformations are used to normalise the variables: the natural logarithm for the two dependent variables and for the number of children, as well as the inverse for the proportion of the working hours and the funding variables.

appropriate for developing countries, which rely on international and philanthropic funding organisations.

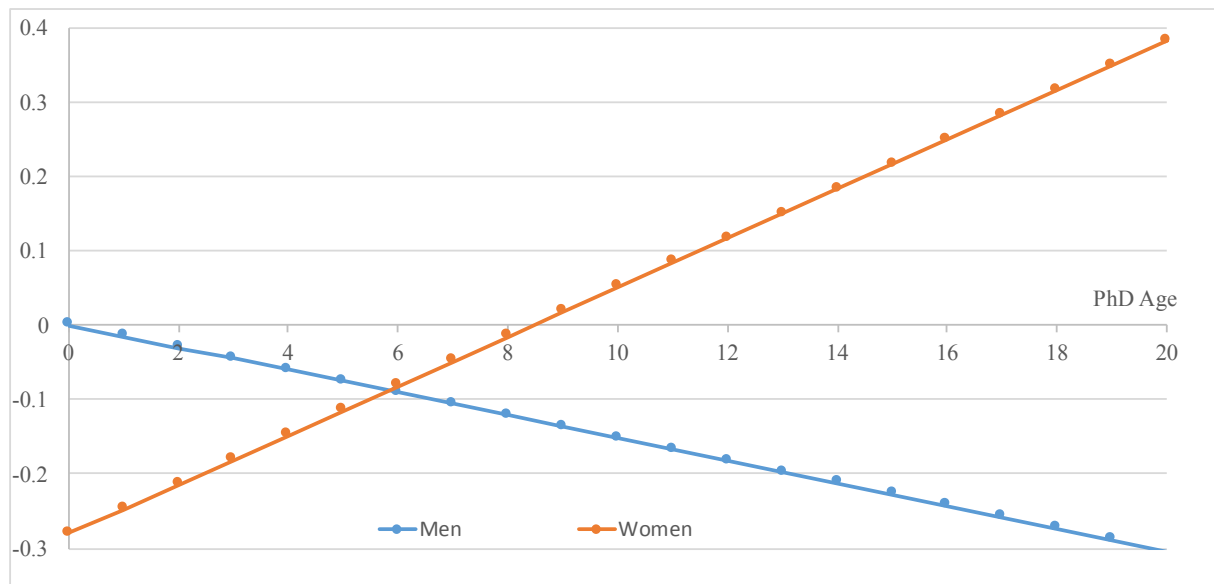


Figure 1. Impact of gender on scientific production (number of papers, book chapters and conference presentations) with respect to ‘PhD’ or career age

Finally, a higher frequency of collaboration is clearly associated with higher research productivity. Because of the importance of international funding, we included foreign collaboration in the regressions. This latter type of collaboration does not impact research productivity on its own but requires moderating effects from various other indicators to have an influence: for instance, the number of hours devoted to research or to fundraising (column Art-8), „PhD“ age – or career maturity – (column Art-11), or even gender (column Art-11). These point towards a more complex framework to be able to fully capture the influence of such an indicator.

DISCUSSION AND CONCLUSIONS

In this paper, we set out to examine four hypotheses corresponding to four types of factors that should have an impact of research production. The first hypothesis, related to researchers’ age, is only very partially supported, which is not surprising considering the fact that our sample is composed mainly of young scientists. Only when interacted with gender and with the proportion of hours dedicated to various tasks is „PhD“ or career age significant.

The hypothesis that female researchers are less prolific is rejected once we account for a variety of factors that influence scientific production. Our hypothesis is only significant when a moderating variable is used. These variables are: „PhD“ or career age, the number of children, the proportion of hours devoted to research or foreign collaboration. For instance, our results clearly show that older women improve their performance as they age.

Table 1. Regression results for the number of articles, chapters and conferences (OLS)

ln(nbArtChapConf+1)	(Art-1)	(Art-2)	(Art-3)	(Art-4)	(Art-5)	(Art-6)
<i>dFemale</i>	-0.0386 (0.0876)	0.0787 (0.1082)		-0.0566 (0.0872)	-0.0481 (0.0874)	-0.0324 (0.0872)
<i>PhDAge</i>	-0.0036 (0.0098)	-0.0035 (0.0097)	-0.0036 (0.0098)	-0.1506 ** (0.0587)	0.1613 * (0.0861)	-0.4452 ** (0.2039)
ln(nbChildren+1)	0.2501 *** (0.0874)	0.3848 *** (0.1139)		0.2292 *** (0.0871)	0.2391 *** (0.0872)	0.2512 *** (0.0869)
1/(propHoursTeach+1)	-0.8918 * (0.4985)	-0.8682 * (0.4968)	-0.8632* (0.5010)	-1.8547 *** (0.6233)	-0.9640* (0.4978)	-0.8696 * (0.4957)
<i>1/(propHoursResearch+1)</i>	-0.0618 (0.4996)	-0.0238 (0.4981)	-0.0674 (0.5000)	0.0941 (0.4991)	0.9436 (0.7205)	0.0238 (0.4982)
<i>1/(propHoursCons+1)</i>	0.1304 (0.5562)	0.1170 (0.5542)	0.1453 (0.5575)	0.2351 (0.5530)	0.0893 (0.5542)	0.1565 (0.5531)
1/(propHoursFund+1)	-0.5100 (0.7920)	-0.4511 (0.7897)	-0.4336 (0.7937)	-0.3456 (0.7879)	-0.6276 (0.7909)	-3.0631 ** (1.4165)
<i>PropSelfHousework</i>	0.0009 (0.0015)	0.0013 (0.0015)	0.0012 (0.0015)	0.0015 (0.0015)	0.0011 (0.0015)	0.0010 (0.0015)
<i>1/(FundNational+1)</i>	-0.1309 (0.1089)	-0.1219 (0.1087)	-0.1159 (0.1097)	-0.1000 (0.1087)	-0.1215 (0.1086)	-0.1004 (0.1092)
1/(FundPrivate+1)	-0.2249 * (0.1198)	-0.2102 * (0.1196)	-0.2227* (0.1200)	-0.2100 * (0.1189)	-0.2120* (0.1194)	-0.2295 * (0.1191)
<i>1/(FundPhil+1)</i>	-0.0669 (0.1730)	-0.0617 (0.1724)	-0.0466 (0.1730)	-0.1264 (0.1732)	-0.0684 (0.1723)	-0.0645 (0.1720)
1/(FundInt+1)	-0.4286 *** (0.1280)	-0.4345 *** (0.1276)	-0.4502 *** (0.1273)	-0.4384 *** (0.1270)	-0.4184 *** (0.1276)	-0.4146 *** (0.1274)
<i>dMobility</i>	0.0182 (0.0951)	0.0019 (0.0951)	-0.0061 (0.0957)	0.0423 (0.0948)	0.0496 (0.0961)	0.0503 (0.0957)
<i>CollForeign</i>	0.0543 (0.0470)	0.0540 (0.0468)	0.0642 (0.0472)	0.0546 (0.0466)	0.0509 (0.0469)	0.0522 (0.0468)
<i>CollNational</i>	0.1717 *** (0.0609)	0.1736 *** (0.0607)	0.1632 *** (0.0611)	0.1714 *** (0.0604)	0.1789 *** (0.0608)	0.1749 *** (0.0606)
dFemale x ln(nbChildren)		-0.2684 * (0.1463)				
Men_with_children			0.3535 *** (0.1360)			
ChildlessWomen			0.0209 (0.1129)			
Women_with_children			0.2261 * (0.1274)			
PhDAge x 1/(PropHoursTeach+1)				0.1951 ** (0.0770)		
PhDAge x 1/(PropHoursResearch+1)					-0.2140* (0.1110)	
PhDAge x 1/(PropHoursFund+1)						0.4639 ** (0.2139)
Constant	3.3279 ** (1.4869)	3.1318 ** (1.4852)	3.1613** (1.5004)	3.6754 ** (1.4807)	2.6969* (1.5163)	5.6558 *** (1.8270)
Country dummy variables	yes	yes	yes	yes	yes	yes
Number of observations	338	338	338	338	338	338
Loglikelihood	-369.97	-368.18	-369.87	-366.57	-367.99	-367.48
F	9.960 ***	9.717 ***	9.470***	9.955 ***	9.744 ***	9.821 ***
R ²	0.3859	0.3924	0.3863	0.3981	0.3930	0.3949
Adjusted R ²	0.3472	0.3520	0.3455	0.3582	0.3527	0.3547

Notes: ***, **, * represent significance at the 1%, 5% and 10% levels respectively; Standard errors in parentheses.

☐

Table 1. (Cont'd)

ln(nbArtChapConf+1)	(Art-7)	(Art-8)	(Art-9)	(Art-10)	(Art-11)	(Art-12)
<i>dFemale</i>	-0.0377 (0.0878)	-0.0394 (0.0870)	-0.0418 (0.0877)	-0.2806** (0.1403)	-0.0188 (0.4048)	-0.0408 (0.0878)
<i>PhDAge</i>	-0.0035 (0.0098)	-0.0038 (0.0097)	-0.0038 (0.0098)	-0.0152 (0.0110)	0.0362 (0.0579)	-0.0034 (0.0098)
dChildren						0.2723*** (0.0988)
ln(nbChildren+1)	0.2506*** (0.0876)	0.2523*** (0.0868)	0.2545*** (0.0876)	0.2367*** (0.0871)	0.2498*** (0.0884)	
1/(propHoursTeach+1)	-0.8895* (0.4993)	-0.8997* (0.4951)	-0.8940* (0.4987)	-0.9542* (0.4963)	-0.9504* (0.5026)	-0.9010* (0.4989)
1/(propHoursResearch+1)	0.1974 (1.0057)	-0.0167 (0.4965)	-0.0813 (0.5003)	0.0337 (0.4984)	-0.0085 (0.5035)	-0.0862 (0.4993)
1/(propHoursCons+1)	0.1227 (0.5576)	0.1159 (0.5524)	0.0975 (0.5578)	0.2128 (0.5541)	0.1440 (0.5602)	0.1333 (0.5571)
1/(propHoursFund+1)	-0.5362 (0.7980)	-5.0942** (2.1293)	-4.1358 (4.3553)	-0.4823 (0.7873)	-0.4973 (0.7942)	-0.4650 (0.7925)
<i>PropSelfHousework</i>	0.0010 (0.0015)	0.0011 (0.0015)	0.0009 (0.0015)	0.0011 (0.0015)	0.0011 (0.0015)	0.0010 (0.0015)
1/(FundNational+1)	-0.1297 (0.1092)	-0.1413 (0.1083)	-0.1288 (0.1090)	-0.1123 (0.1086)	-0.1309 (0.1092)	-0.1244 (0.1092)
1/(FundPrivate+1)	-0.2256* (0.1200)	-0.2508** (0.1195)	-0.2291* (0.1199)	-0.2271* (0.1191)	-0.2288* (0.1208)	-0.2282* (0.1198)
1/(FundPhi+1)	-0.0681 (0.1733)	-0.0924 (0.1722)	-0.0536 (0.1738)	-0.0938 (0.1724)	-0.0894 (0.1739)	-0.0497 (0.1729)
1/(FundIntl+1)	-0.4333*** (0.1291)	-0.4120*** (0.1273)	-0.4130*** (0.1294)	-0.4223*** (0.1273)	-0.4477*** (0.1281)	-0.4509*** (0.1272)
<i>dMobility</i>	0.0188 (0.0952)	0.0212 (0.0944)	0.0148 (0.0952)	0.0598 (0.0964)	0.0369 (0.0957)	0.0040 (0.0949)
CollForeign	0.1324 (0.2672)	-1.6375** (0.7317)	0.0554 (0.0471)	0.0543 (0.0467)	0.1507* (0.0883)	0.0623 (0.0471)
CollNational	0.1703*** (0.0612)	0.1723*** (0.0605)	-0.7713 (1.1154)	0.1736*** (0.0606)	0.1469 (0.1123)	0.1651*** (0.0610)
<i>CollForeign</i> x 1/(PropHoursResearch+1)	-0.1111 (0.3741)					
CollForeign x 1/(PropHoursFund+1)		1.7453** (0.7533)				
<i>CollNational</i> x 1/(PropHoursFund+1)			0.9636 (1.1382)			
dFemale x <i>PhDAge</i>				0.0483** (0.0220)		
<i>dFemale</i> x <i>CollForeign</i>					0.0535 (0.0868)	
<i>dFemale</i> x <i>CollNational</i>					-0.0440 (0.1177)	
PhDAge x <i>CollForeign</i>					-0.0254** (0.0118)	
<i>PhDAge</i> x <i>CollNational</i>					0.0091 (0.0155)	
Constant	3.1801** (1.5700)	7.7838*** (2.4249)	6.9253 (4.5021)	3.2416** (1.4785)	3.1425** (1.5483)	3.3061** (1.4905)
Country dummy variables	yes	yes	yes	yes	yes	yes
Number of observations	338	338	338	338	338	338
Loglikelihood	-369.92	-367.12	-369.59	-367.40	-367.08	-370.28
F	9.463***	9.872***	9.512***	9.831***	8.562***	9.914***
R ²	0.3860	0.3962	0.3873	0.3952	0.3963	0.3848
Adjusted R ²	0.3453	0.3560	0.3466	0.3550	0.3500	0.3460

Notes: ***, **, * represent significance at the 1%, 5% and 10% levels respectively; Standard errors in parentheses.

Table 2. Regression results for the number of patents (Tobit)

$\ln(\text{nbPatents}+1)$	(Pat-1)	(Pat-2)	(Pat-3)	(Pat-4)	(Pat-5)	(Pat-6)	(Pat-7)
<i>dFemale</i>	-0.0905 (0.1176)	-0.0378 (0.1480)	-1.5940 ** (0.7953)		-0.0927 (0.1179)	-0.0938 (0.1177)	-0.0895 (0.1177)
<i>PhDAge</i>	-0.0082 (0.0142)	-0.0080 (0.0142)	-0.0071 (0.0141)	-0.0050 (0.0141)	-0.0333 (0.0893)	0.0633 (0.1214)	-0.0678 (0.2861)
$\ln(\text{nbChildren}+1)$	0.0026 (0.1171)	0.0612 (0.1541)	0.0016 (0.1167)	0.0327 (0.1156)	0.0005 (0.1173)	0.0001 (0.1172)	0.0035 (0.1172)
1/(propHoursTeach+1)	-1.5539 ** (0.6805)	-1.5431 ** (0.6804)	-1.5113 ** (0.6783)	-1.6365 ** (0.6706)	-1.7091 * (0.8704)	-1.5777 ** (0.6814)	-1.5510 ** (0.6806)
$1/(\text{propHoursResearch}+1)$	0.3290 (0.6871)	0.3430 (0.6874)	-0.6256 (0.8489)	0.2598 (0.6766)	0.3526 (0.6919)	0.7691 (1.0112)	0.3392 (0.6889)
$1/(\text{propHoursCons}+1)$	1.1043 (0.7877)	1.0966 (0.7875)	1.0613 (0.7886)	1.0808 (0.7798)	1.1230 (0.7904)	1.0920 (0.7880)	1.1059 (0.7877)
$1/(\text{propHoursFund}+1)$	-0.3549 (1.0931)	-0.3317 (1.0943)	-0.3001 (1.0914)	-0.1396 (1.0833)	-0.3358 (1.0947)	-0.3981 (1.0956)	-0.6848 (1.9223)
<i>PropSelfHousework</i>	0.0007 (0.0021)	0.0009 (0.0021)	0.0011 (0.0021)	0.0007 (0.0020)	0.0008 (0.0021)	0.0008 (0.0021)	0.0007 (0.0021)
$1/(\text{FundNational}+1)$	-0.1123 (0.1478)	-0.1074 (0.1480)	-0.1042 (0.1473)	-0.1200 (0.1455)	-0.1083 (0.1485)	-0.1094 (0.1479)	-0.1083 (0.1490)
$1/(\text{FundPrivate}+1)$	-0.2026 (0.1612)	-0.1968 (0.1615)	-0.2265 (0.1613)	-0.2138 (0.1589)	-0.2004 (0.1614)	-0.1983 (0.1613)	-0.2033 (0.1613)
$1/(\text{FundPhil}+1)$	0.0969 (0.2398)	0.0967 (0.2397)	0.0752 (0.2397)	0.1630 (0.2384)	0.0887 (0.2415)	0.0954 (0.2399)	0.0967 (0.2399)
1/(FundInt+1)	-0.3458 ** (0.1714)	-0.3497 ** (0.1715)	-0.3372 ** (0.1707)	-0.4005 ** (0.1697)	-0.3471 ** (0.1714)	-0.3398 ** (0.1716)	-0.3427 ** (0.1720)
<i>dMobility</i>	0.0068 (0.1262)	-0.0003 (0.1267)	-0.0059 (0.1260)		0.0096 (0.1266)	0.0181 (0.1276)	0.0109 (0.1277)
<i>CollForeign</i>	-0.0158 (0.0629)	-0.0159 (0.0629)	-0.0122 (0.0627)	-0.0231 (0.0619)	-0.0157 (0.0629)	-0.0172 (0.0629)	-0.0161 (0.0629)
CollNational	0.1955 ** (0.0837)	0.1966 ** (0.0837)	0.1835 ** (0.0836)	0.2091 ** (0.0825)	0.1954 ** (0.0837)	0.1991 ** (0.0840)	0.1966 ** (0.0839)
<i>dFemale</i> x $\ln(\text{nbChildren})$		-0.1151 (0.1964)					
dFemale x $1/(\text{PropHoursResearch}+1)$			2.0674 * (1.0812)				
MobileMen				-0.4139 ** (0.1776)			
NonMobileWomen				-0.6204 *** (0.1974)			
<i>MobileWomen</i>				-0.2397 (0.1772)			
<i>PhDAge</i> x $1/(\text{PropHoursTeach}+1)$					0.0326 (0.1141)		
<i>PhDAge</i> x $1/(\text{PropHoursResearch}+1)$						-0.0939 (0.1586)	
<i>PhDAge</i> x $1/(\text{PropHoursFund}+1)$							0.0623 (0.2988)
Constant	0.6790 (2.0277)	0.5986 (2.0322)	1.3738 (2.0534)	0.9273 (1.9992)	0.7431 (2.0399)	0.3863 (2.0858)	0.9804 (2.4897)
Country dummy variables	yes	yes	yes	yes	yes	yes	yes
sigma	0.9455 *** (0.0494)	0.9450 *** (0.0493)	0.9411 *** (0.0491)	0.9292 *** (0.0484)	0.9453 *** (0.0493)	0.9450 *** (0.0493)	0.9454 *** (0.0494)
Number of observations	338	338	338	338	338	338	338
Loglikelihood	-392.77	-392.60	-390.94	-387.26	-392.73	-392.59	-392.75
χ^2	133.833 ***	134.176 ***	137.5 ***	144.85 ***	133.915 ***	134.186 ***	133.876 ***
Pseudo R ²	0.1456	0.1459	0.1496	0.1575	0.1457	0.1460	0.1456

Notes: ***, **, * represent significance at the 1%, 5% and 10% levels respectively; Standard errors in parentheses. Number of left-censored observations = 118.

Table 2. (Cont'd)

ln(nbPatents+1)	(Pat-8)	(Pat-9)	(Pat-10)	(Pat-11)	(Pat-12)	(Pat-13)	(Pat-14)
<i>dFemale</i>	-0.0853 (0.1165)	-0.0910 (0.1168)	-0.0985 (0.1169)	0.0686 (0.1906)	-0.1910 (0.5505)	-0.0911 (0.1177)	
<i>PhDAge</i>	-0.0069 (0.0141)	-0.0085 (0.0141)	-0.0078 (0.0140)	0.0002 (0.0158)	-0.1880 ** (0.0772)	-0.0084 (0.0142)	-0.0084 (0.0142)
<i>dChildren</i>						0.0126 (0.1344)	
ln(nbChildren+1)	0.0054 (0.1161)	0.0060 (0.1164)	0.0190 (0.1168)	0.0117 (0.1173)	0.0093 (0.1156)		
1/(propHoursTeach+1)	-1.5121 ** (0.6734)	-1.5674 ** (0.6756)	-1.5541 ** (0.6759)	-1.5255 ** (0.6798)	-1.1975 * (0.6691)	-1.5459 ** (0.6813)	-1.5093 ** (0.6837)
1/(propHoursResearch+1)	3.1516 ** (1.3879)	0.4268 (0.6836)	0.2773 (0.6832)	0.2638 (0.6887)	0.4761 (0.6734)	0.3338 (0.6862)	0.3484 (0.6865)
1/(propHoursCons+1)	1.0087 (0.7782)	1.0900 (0.7818)	0.9960 (0.7825)	1.0427 (0.7883)	1.3935 * (0.7703)	1.1119 (0.7886)	1.1227 (0.7887)
1/(propHoursFund+1)	-0.6233 (1.0891)	-6.7196 ** (2.9938)	-11.4000 * (6.3234)	-0.3896 (1.0919)	-0.0551 (1.0724)	-0.3541 (1.0929)	-0.3277 (1.0944)
<i>PropSelfHousework</i>	0.0016 (0.0021)	0.0010 (0.0021)	0.0006 (0.0021)	0.0007 (0.0021)	0.0007 (0.0020)	0.0008 (0.0021)	0.0010 (0.0021)
1/(FundNational+1)	-0.0993 (0.1466)	-0.1237 (0.1470)	-0.1029 (0.1470)	-0.1208 (0.1477)	-0.0673 (0.1446)	-0.1115 (0.1480)	-0.1023 (0.1487)
1/(FundPrivate+1)	-0.2157 (0.1598)	-0.2372 (0.1609)	-0.2079 (0.1603)	-0.2007 (0.1611)	-0.1816 (0.1582)	-0.2018 (0.1612)	-0.1978 (0.1612)
1/(FundPhil+1)	0.0880 (0.2379)	0.0513 (0.2391)	0.1564 (0.2408)	0.1132 (0.2399)	0.1299 (0.2358)	0.0962 (0.2393)	0.0971 (0.2393)
1/(FundInt+1)	-0.3973 ** (0.1714)	-0.3229 * (0.1707)	-0.3071 * (0.1716)	-0.3508 ** (0.1712)	-0.3154 * (0.1671)	-0.3447 ** (0.1703)	-0.3455 ** (0.1702)
<i>dMobility</i>	0.0142 (0.1252)	0.0130 (0.1253)	0.0006 (0.1255)	-0.0161 (0.1279)	-0.0346 (0.1236)	0.0067 (0.1259)	-0.0027 (0.1269)
CollForeign	0.8268 ** (0.3668)	-2.4212 ** (1.0636)	-0.0158 (0.0626)	-0.0148 (0.0628)	-0.4056 *** (0.1179)	-0.0155 (0.0630)	-0.0138 (0.0630)
CollNational	0.1806 ** (0.0830)	0.1945 ** (0.0833)	-2.6565 * (1.6082)	0.1920 ** (0.0836)	0.2348 (0.1492)	0.1954 ** (0.0837)	0.1939 ** (0.0837)
CollForeign x 1/(PropHoursResearch+1)	-1.1926 ** (0.5115)						
CollForeign x 1/(PropHoursFund+1)		2.4802 ** (1.0940)					
CollNational x 1/(PropHoursFund+1)			2.9110 * (1.6395)				
<i>dFemale x PhDAge</i>				-0.0322 (0.0304)			
dFemale x CollForeign					0.3010 *** (0.1147)		
<i>dFemale x CollNational</i>					-0.1756 (0.1566)		
PhDAge x CollForeign					0.0498 *** (0.0159)		
<i>PhDAge x CollNational</i>					0.0131 (0.0209)		
<i>Men_with_children</i>							0.0874 (0.1852)
<i>ChildlessWomen</i>							-0.0329 (0.1539)
<i>Women_with_children</i>							-0.0803 (0.1724)
Constant	-0.9576 (2.1260)	6.8228 ** (3.3597)	11.6477 * (6.4887)	0.7685 (2.0256)	0.5929 (2.0553)	0.6535 (2.0336)	0.5228 (2.0456)
Country dummy variables	yes	yes	yes	yes	yes	yes	yes

sigma	0.9362 *** (0.0488)	0.9385 *** (0.0490)	0.9394 *** (0.0490)	0.9439 *** (0.0493)	0.9189 *** (0.0479)	0.9454 *** (0.0494)	0.9450 *** (0.0493)
Number of observations	338	338	338	338	338	338	338
Loglikelihood	-390.06	-390.08	-391.19	-392.21	-384.71	-392.77	-392.60
c2	139.244 ***	139.210 ***	137.002 ***	134.950 ***	149.958 ***	133.841 ***	134.185 ***
Pseudo R ²	0.1515	0.1514	0.1490	0.1468	0.1631	0.1456	0.1460

Notes: ***, **, * represent significance at the 1%, 5% and 10% levels respectively; Standard errors in parentheses. Number of left-censored observations = 118.

□

Our third hypothesis on funding variables only highlights the importance of private funding and of international funding for research output. Hypothesis H3a cannot be validated. Private and international funding matter more for classic research output. Surprisingly, private funding has no impact on technological output. We therefore cannot validate hypothesis H3b, nor H3c. For all four categories of output, however, international funding has a strong positive relationship with scientific output.

The last hypothesis is the only one that is wholeheartedly supported, hence validating the close relationship between collaboration and research output of any kind. A higher frequency of collaboration is clearly associated with higher research productivity. This goes beyond the scope of this paper but is a very promising avenue for research.

This research is based on a single survey on the perception of researchers about their career and research outputs. As such, there are a number of limitations to this study. First, out of the 750 responses, only 338 are usable for our regression analysis. We soon realised that the questionnaire was too long. This will have to be remedied in future similar studies. Second, the survey was entirely anonymous and as a consequence, we cannot verify the true output of these researchers using a standard bibliometric tool, but more importantly, the survey cannot be used to further study these researchers in the future to see whether their perceptions will have had an impact on their future career.

In terms of policy, the take-home message from this paper is clearly that the importance of foreign funding has an influence that is not noticed in developed countries that have well-developed grant-awarding organisations. Foreign collaboration and mobility also have a more complex impact that needs to be further investigated. These foreign relations are important and may compensate deficiencies in the local science system.

REFERENCES

- Abramo, G., D'Angelo, C.A., Caprasecca, A. (2009a). Gender differences in research productivity: A bibliometric analysis of the Italian academic system. *Scientometrics* 79, 517-539.
- Beaudry, C., Allaoui, S. (2012). Impact of public and private research funding on scientific production: The case of nanotechnology. *Research Policy*, 41(9), 1589-1606.
- Daniel, H.Z., Hempel, D.J., Srinivasan, N. (2003). Project selection: A process analysis. *Industrial Marketing Management*, 32, 39-54.
- DesRoches, C.M., Zinner, D.E., Sowmya, R.R., Iezzoni, L.I., Campbell, E.G. (2010). Activities, productivity, and compensation of men and women in the life sciences. *Academic Medicine*, 85(4), 631-639.

- Dietrich, A., Srinivasan, N. (2007). The optimal age to start a revolution. *Journal of Creative Behavior*, 41(1), 339-351.
- Feist, G.J. (2006). *The psychology of science and the origins of the scientific mind*. New Haven: Yale University Press.
- Fox, M.F. (1991). Gender, environmental milieu, and productivity in science. In Zuckerman H., Cole J.R., Bruer J.T. (eds). *The Outer Circle. Women in the Scientific Community*. New York: W.W. Norton & Company. pp. 188-204.
- Fox, M.F. (2005). Gender, family characteristics, and publication productivity among scientists. *Social Studies of Science*, 35, 131-150.
- Jacob, B., Lefgren, L. (2007). The impact of research grant funding on scientific productivity. NBER Working Paper No. 13519.
- Kyvik, S., Teigen, M. (1996). Child care, research collaboration, and gender differences in scientific productivity. *Science, Technology & Human Values* 21, 54.
- Larivière, V., Ni, C., Gingras Y., Cronin, B., Sugimoto, C. (2013). Global gender disparities in science. *Nature*, 12 (December), 211-213.
- Leahey, E. (2006). Gender differences in productivity. *Gender & Society*, 20, 754-780.
- Lehman, H.C. (1953). *Age and Achievement*. Princeton: Princeton University Press.
- Long, J.S. (1992). Measures of sex differences in scientific productivity. *Social Forces*, 71(1), 159-178.
- Manjarrés-Henríquez, L., Gutiérrez-Gracia, A., Carrión-García, A., Vega-Jurado, J. (2009). The effects of university–industry relationships and academic research on scientific performance: Synergy or substitution? *Research in Higher Education*, 50(8), 795-811.
- MIT (1999). *A study on the status of women faculty in science at MIT*. Cambridge: Massachusetts Institute of Technology.
- Over, R. (1988). Does scholarly impact decline with age? *Scientometrics*, 13(5–6), 215–223.
- Zucker, L.G., Darby, M.R., Furner, J., Liu, R.C., Ma, H. (2007). Minerva unbound: Knowledge stocks, knowledge flows and new knowledge production. *Research Policy*, 36, 850-863.

Appendix

Table A1 – Variable description and gender mean-comparison tests

Variable	Description	Men	Women	M-W^a
nbArtChapConf	Number of peer-reviewed articles, book chapters and presentations at conferences	19.952	16.596	0.6174
nbPatents	Number of pending and granted patents	2.560	2.467	0.8534
dFemale	Dummy variable that takes the value 1 for women and 0 for men			
PhDAge	Number of years since PhD graduation (2015 – PhD graduation)	4.943	4.670	0.8023
dChildren	Dummy variable taking the value 1 if the researcher has children between the ages of 0 and 18	0.3956	0.3852	0.8037
nbChildren	Number of children between the ages of 0 and 18	0.7802	0.7185	0.8074
dMalaysia	Dummy variable taking the value 1 if the researcher is located in Malaysia	0.2821	0.3407	0.1401
dThailand	Dummy variable taking the value 1 if the researcher is located in Thailand	0.3773	0.5000	0.0040
propHoursTeach	Proportion of working hours dedicated to teaching	6.837	8.537	0.0243
propHoursResearch	Proportion of working hours dedicated to research, training and supervision	19.603	21.577	0.0741
propHoursCons	Proportion of working hours dedicated to consulting or implementing researcher	4.837	3.632	0.0420
propHoursFund	Proportion of working hours dedicated to fundraising	2.063	1.211	0.0084
PropSelfHousework	Proportion of the housework performed by the researcher	50.392	55.905	0.0342
FundNational	Proportion of a researchers funding that comes from National sources (local and national governments)	35.000	34.998	0.9571
FundPrivate	Proportion of a researchers funding that comes from private sources (business firms or industry)	4.912	3.299	0.0052
FundPhil	Proportion of a researchers funding that comes from philanthropic sources (private not-for profit foundations/agencies)	2.442	1.944	0.4131
FundInt	Proportion of a researchers funding that comes from international organisations	4.716	4.578	0.4003
dMobility	Dummy variable that takes the value 1 if the researcher has lived, studied or worked for more than 3 months in a country other than his/her home country in the past 5 years	0.733	0.659	0.0634
CollForeign ^b	Average value of the importance of collaborating on publications and on research projects with foreign partners	2.525	2.362	0.0858
CollNational ^b	Average value of the importance of collaborating on publications and on research projects with partners from the same country	3.558	3.601	0.5776

Notes: ^a Significance of the Mann-Whitney two-sample statistic to compare two populations; ^b These two variables are the result from a principal component analysis presented in Table A2.

□

Table A2 – Principal component analysis results for reducing the collaboration dimensions

Items	Components		
	CollForeign	CollOthDiscGender	CollNational
How often does the researcher collaborates on publications with researchers from:			
Own university / research organization	-0.034	0.350	0.710
Other universities or research organizations in own country	0.454	0.071	0.645
Other countries (same continent)	0.810	0.054	0.178
Other countries (other continent)	0.753	0.362	0.075
Other disciplines / research fields	0.310	0.591	0.307
Other gender	0.177	0.763	0.245
Private companies		Eliminated	
How often does the researcher collaborates on research projects with researchers from:			
Own university / organization	-0.089	0.381	0.716
Other organizations in your own country	0.391	0.112	0.707
Other countries (same continent)	0.836	0.139	0.146
Other countries (other continent)	0.777	0.395	-0.009
Other disciplines/ research fields	0.270	0.700	0.232
Other gender	0.136	0.839	0.149
Kaiser-Mayer-Olkin (K-M-O) index	0.752		
Eigenvalues	3.113	2.720	2.222
% Variance	25.946	22.667	18.520
% Variance Cumulative	25.946	48.613	67.133
Cronbach's alpha	0.873	0.828	0.725

Notes: All collaboration items are measured by a five-point Likert scale (1 = Never or very rarely; to 5 = Very often or always).

☐



Identifying the Gender Dimension in Research Content¹

David Campbell^{*}, Bastien St. Louis Lalonde^{*}, Chantale Tippett^{*}, Éric Archambault^{*}, Julie Callaert^{**}, Katerina Mantouvalou^{***}, and Lucy Arora^{***}

^{*}*david.campbell@science-metrix.com; bastien.st-louislalonde@science-metrix.com; chantale.tippett@science-metrix.com; eric.archambault@science-metrix.com;*

***julie.callaert@kuleuven.be;*

****katerina.mantouvalou@icfi.com; lucy.arora@icfi.com*

Science-Metrix, 1335 Avenue du Mont-Royal Est, Montréal, H2J 1Y6 (Canada)

KU Leuven, ECOOM – INCENTIM - Faculty of Business and Economics - Department MSI

Waaistraat 6 – Bus 3536, 3000 Leuven (Belgium)

ICF International, Watling House, 33 Cannon Street, London, EC4M 5SB (UK)

ABSTRACT

Globally, there is an increasing interest in integrating the gender dimension in research content (GDRC). As a first step towards monitoring progress in this area, a new indicator measuring the proportion of a country's scientific publications integrating a gender dimension in their subject matter was developed for the European Commission's *She Figures 2015* publication. This indicator is based on a keyword-based query covering both sex-related terms (biological characteristics of both women and men) and gender-related terms (social/cultural factors of both women and men). The final GDRC dataset consisted of some 212,600 distinct publications including a gender dimension in their research content. Findings suggest that integrating a gender dimension into research content is relatively rare. Unsurprisingly, it was less common for scientific articles in the fields of agricultural sciences, engineering and technology, and natural sciences to do so, and more common in the social sciences.

Keywords

Gender dimension in research content; sex/gender analysis; science policy

INTRODUCTION

Within the context of the European Commission's Eighth Framework Programme for Research and Technological Development (Horizon 2020), activities towards achieving gender equality are being implemented along three main axes: fostering gender balance in research teams; ensuring gender balance in decision-making; and integrating gender analysis in research and innovation (R&I) content (European Commission & Directorate-General for

¹ The production of this methodology for the *She Figures 2015* was led by Science-Metrix, with extensive contributions from the Directorate-General for Research and Innovation of the European Commission; EIGE; Eurostat; the Helsinki Group on Gender in Research and Innovation and their Statistical Correspondents; ICF International, KU Leuven and the OECD.

Research and Innovation, 2014). Where relevant, Horizon 2020 participants must now specify in their grant proposals how the gender dimension (i.e. taking into account as relevant the biological characteristics and the social and cultural features of both women and men) will be integrated into the subject matter of their projects (European Parliament and the Council of the European Union, 2013). As such, it is highly relevant to begin monitoring trends in the extent to which researchers incorporate such aspects into their research content. This will provide evidence to allow the monitoring of the extent to which the gender dimension appears in the research content of the scientific outputs produced by countries.

In the context of the production of the *She Figures 2015* publication (European Commission, 2016), a new indicator was developed to monitor the extent to which researchers integrate a gender dimension into their research content. Some experts involved in the consultation process (acknowledged at the beginning of this paper) expressed concerns about the scope of the topics that should be included in the gender dimension, which led to a consideration of the definitions of ‘sex’, ‘gender’, ‘sex/gender analysis’, and ‘gender dimension in research’ in the glossary of the document *Gender Equality in Horizon 2020* (European Commission & Directorate-General for Research and Innovation, 2014).

METHODS

The following sections provide an overview of some of the key methodological considerations in the development of this indicator; however, for a more complete summary, the reader is referred to the comprehensive methodological document produced for the *She Figures 2015* study (Campbell, 2015).

Bibliographic data on peer-reviewed scientific publications is necessary for producing data on this topic. Thomson Reuters’ Web of Science (WoS) was selected for this study because it provides cited references for each document it includes (e.g. articles or chapters published in a journal or book series), as well the institutional affiliations of authors. The two-step process to construct the query for the retrieval of scientific publications in which a gender dimension is addressed comprises

1. the identification of an initial set (i.e. ‘seed’) of highly relevant papers, and
2. the extraction of gender-specific terms through an analysis of the textual content present in the seed, which are then used to expand the seed to obtain the final dataset.

Based on the definitions provided in the Horizon 2020 documentation cited above, the gender dimension in research content includes both the concepts of sex and gender as well as the concept of sex/gender analysis in humans.

In addition to research outputs focused on a well-defined gender topic (e.g. feminism, gender pay gap, gender equality), research content in which a distinction or a comparison is made between men and women either in the title, abstract, or author keywords of scientific publications were deemed relevant. Following extensive consultation with the *She Figures 2015* expert committee, research outputs studying the animal kingdom (e.g. feminisation of fish populations) as well as other non-human biological entities (e.g. plants), were excluded in the construction of the dataset on the gender dimension in research content. Moreover, scientific papers investigating medical conditions specific to one gender (e.g. menopause, erectile dysfunction) were also not to be considered pertinent to the dataset as the inclusion of

those would result in the inclusion of a very large portion of scientific publications in the medical fields.

Creation of the seed dataset

All publications indexed in the WoS were classified into six large domains (Applied Sciences, Arts & Humanities, Economic & Social Sciences, General, Health Sciences and Natural Sciences), then further divided into 22 fields and 176 subfields using Science-Metrix's journal-based classification (Archambault, Beauchesne, & Caruso, 2011). This classification is mutually exclusive – that is, each article is classified into one and only one set of domain, field and subfield. Using the fields of science and technology (FOS) classification in the *Frascati Manual* (OECD, 2002), as well as the revised classification (OECD, 2007), the subfields in Science-Metrix's classification were then matched to their corresponding FOS as defined in the *Frascati Manual's* 2007 description. Thus, this indicator was computed for each of the following six FOS:

- (NS) Natural sciences;
- (ET) Engineering and Technology;
- (MS) Medical sciences;
- (AS) Agricultural sciences;
- (SS) Social sciences; and
- (H) Humanities.

The first step in identifying scientific publications relevant to the gender dimension in research content was to identify fields and subfields directly related to gender research. The subfield of Gender Studies under the field of Social Sciences was found to be directly relevant, and contained 6,023 publications discussing a gender-related topic (for the period 2002–2013). A validation check (title and abstract reading) of a randomly selected sample of 100 articles was performed to enable the confirmation of the pertinence of the publications in this subfield.

In the subsequent step, a search for journal names containing the term 'gender' was executed in the WoS. The scope of each journal was evaluated either by accessing its website or, when this was not possible, by examining the journal's publications (Table 1). Based on the scope of the journals, a verdict determining if the papers published in each journal were pertinent to the gender dimension in research content was assigned. All the publications contained in the relevant journals (2,150 for 2002–2013) were then added to the seed dataset.

Table 1. WoS journals containing the term *gender* in their name

Journal name	Papers (2002–2013)	Verdict
FOCUS ON GENDER: PARENT AND CHILD CONTRIBUTIONS TO THE SOCIALIZATION OF EMOTIONAL COMPETENCE	7	ok
PEDIATRIC GENDER ASSIGNMENT: A CRITICAL REAPPRAISAL	19	ok
GENDER AND LANGUAGE	54	ok
POLITICS & GENDER	92	ok
JOURNAL OF WOMENS HEALTH & GENDER-BASED MEDICINE	121	ok
INDIAN JOURNAL OF GENDER STUDIES	147	ok
JOURNAL OF GENDER STUDIES	234	ok
GENDER MEDICINE	266	ok
GENDER PLACE AND CULTURE	333	NO (also on race, ethnicity, sexuality, etc)
GENDER WORK AND ORGANIZATION	346	ok
GENDER & SOCIETY	394	ok
GENDER AND EDUCATION	468	ok

Source: Compiled by Science-Metrix from WoS data (Thomson Reuters)

Next, journals that published articles classified in the subfield of Gender Studies, and which were different from those identified previously, were retrieved. The scope of these journals was assessed in the same manner as previously described in order to evaluate the relevance of their content. The 3,700 articles published in appropriate journals between 2002 and 2013 were subsequently added to the seed dataset (Table 2).

Table 2. WoS journals containing publications classified under the subfield Gender Studies

Journal name	Papers (2002–2013)	Verdict
FEMINISTISCHE STUDIEN	101	ok
FEMINIST THEORY	101	ok
NOUVELLES QUESTIONS FEMINISTES	141	ok
INTERNATIONAL FEMINIST JOURNAL OF POLITICS	163	ok
ASIAN JOURNAL OF WOMENS STUDIES	163	ok
HYPATIA-A JOURNAL OF FEMINIST PHILOSOPHY	177	ok
FRONTIERS-A JOURNAL OF WOMEN STUDIES	179	ok
FEMINIST STUDIES	213	ok
MEN AND MASCULINITIES	215	ok
FEMINIST REVIEW	230	ok
SEXUALITIES	232	ok
AUSTRALIAN FEMINIST STUDIES	237	ok
EUROPEAN JOURNAL OF WOMENS STUDIES	256	ok
FEMINISM & PSYCHOLOGY	408	ok
SIGNS	504	ok
JOURNAL OF SEX RESEARCH	510	ok
WOMENS STUDIES INTERNATIONAL FORUM	604	ok

Source: Compiled by Science-Metrix from WoS data (Thomson Reuters)

The last phase in building the seed dataset involved using Medline's controlled vocabulary medical subject heading (MeSH) terms to identify gender-related scientific articles indexed in the WoS and in Medline. A search for MeSH terms containing every variant of gender, *femin**, *women* and *men* allowed for the identification of 18 relevant MeSH terms (Table 3). To enable the extraction of publications using MeSH terms in the WoS, all publications in Medline were matched to their corresponding entry in the WoS. The pertinence of the papers

retrieved by each MeSH term was evaluated through a title and abstract analysis on a random sample of papers. The only MeSH terms that were not satisfactory were Feminization and Pregnant Women.

Table 3. MeSH terms, the number of papers associated with them in the WoS and the verdict for addition of the associated publications in the seed dataset

Mesh descriptor term	Papers (2002-2013)	Verdict
Feminine Hygiene Products	9	NO
Sexual and Gender Disorders	17	ok
Dentists, Women	29	ok
Femininity	51	ok
Feminization	68	NO, medical explanation of feminization (most of the time not in human)
Transgendered Persons	113	ok
Feminism	207	ok
Physicians, Women	323	ok
Men's health	390	ok
Women's Rights	403	ok
Women's Health Services	573	ok
Women, Working	627	ok
Pregnant Women	694	NO, as decided by expert committee
Battered Women	991	ok
Women	1,680	ok
Gender Identity	2,045	ok
Women's Health	4,480	ok

Source: Compiled by Science-Metrix from WoS data (Thomson Reuters)

Following the addition of the articles captured by the selected MeSH terms, the seed dataset consisted of 17,900 distinct publications. A publication may have been retrieved multiple times by the different techniques (subfield, journals or MeSH terms) but was only counted once.

Creation of the final dataset

In this phase, the seed dataset was expanded using a query searching for gender-related terminology in the title, abstract and author keywords of the publications indexed in the WoS. Highly relevant terms were identified using the term frequency–inverse document frequency (TF-IDF) statistic. The TF-IDF statistic determines the importance of a given expression (a term or set of terms) in a specific set of documents (i.e. the seed dataset) relative to a reference collection of documents (the WoS). The relevance of an expression increases proportionally to the number of times it appears in the seed dataset but is offset by the frequency of the word in the reference collection. This operation increases the detection of rare and specific expressions. Two lists of expressions were tested using the TF-IDF weight. The first list consists of the EIGE draft thesaurus on gender equality terms (which contains more than 600 terms). The second list consists of a set of about 10 million noun phrases (i.e. scientific expressions) extracted from the titles, abstracts and author keywords of publications in the WoS.

The keywords of the highest relevance (i.e. those with the highest TF-IDF weight, some of which are shown in Table 4) are the most promising expressions to identify gender-related publications. As one goes down the list, there is a threshold (not necessarily well defined) at which point the expressions are not specific enough to be used in the search query aimed at expanding the seed dataset (highlighted in red).

Table 4. Items with low TF-IDF weight not included in the keyword query

EIGE keywords		WoS extracted noun phrases	
Keyword	Weight	Keyword	Weight
Transsexual	12.59	Intimate partner	14.89
•	•	•	•
•	•	•	•
•	•	•	•
Demographic change	7.62	Risk taking	10.22
Witness	7.62	Awakening response	10.22
Spending	7.60	Problematise	10.22
Career planning	7.59	Source care	10.22
Migrant worker	7.54	Meaning practice	10.22
Families	7.52	Low libido	10.22
Risk group	7.52	Qualitative method	10.22
•	•	•	•
•	•	•	•

Source: Compiled by Science-Metrix from WoS data (Thomson Reuters)

Each of the promising expressions was tested manually by retrieving publications in which the expression appeared in their titles, abstracts or author keywords (using wildcards to ensure all wording with the same root would be retrieved). For each search term, a seasoned analyst read the title and abstract of a random sample of the retrieved publications to judge their pertinence. When the majority of the retrieved publications were judged relevant, the search expression was included in the keyword-based query. The experts group expressed concerns that the approach using the TF-IDF weight could induce a bias towards the seed dataset (Gender Studies subfield + specialist journals + MeSH terms), so the selection of relevant search expressions was re-iterated using the TF-IDF weight, resulting in around 220 search expressions being included in the query. The final step in the query consisted of deleting articles about the animal kingdom that were not filtered by the field/subfield exclusion criteria.

Recall and precision

The dataset's quality was assessed via two parameters often used in information retrieval. The recall (i.e. the percentage of false negatives, or relevant papers that were not retrieved) of the seed dataset was measured by taking the intersection of the publications in the seed dataset with those retrieved by the keyword-based query over the size of the seed dataset (i.e. the percentage of the seed retrieved with the keyword-based query). This facilitates the assessment of how well selected search expressions capture the core literature related to Gender Studies and gender-related MeSH terms. The recall was also assessed at the journal

level by measuring the percentage of articles captured by the keyword query that are present in the journals that were included in the seed dataset.

The recall of the seed dataset is near 60% (Table 5), which is around what is expected for a dataset in the Social Sciences and Humanities (SSH). Indeed, it is difficult to achieve higher recall for subjects related to the SSH, as the expressions used in this domain are usually less specific to a particular area of research than the expressions used in the Natural Sciences and Engineering or Health Sciences domains. Adding more keywords (likely more generic terms) to increase the recall would be detrimental to the dataset, as this would lead to the concomitant retrieval of false positives and a subsequent reduction in precision. The recall of the specialist journals is also satisfactory, with the majority of them having a recall above the 60% mark.

Table 5. Recall of the seed dataset (i.e. gender studies subfield, specialist journals and MeSH terms) and of each of the specialist journals using the keyword-based query

Dataset/Journal	Articles from KW query	Total articles	Recall
Seed dataset	10,453	17,885	58.4%
GENDER & SOCIETY	375	394	95.2%
MEN AND MASCULINITIES	202	215	94.0%
GENDER AND LANGUAGE	50	54	92.6%
JOURNAL OF GENDER STUDIES	215	234	91.9%
EUROPEAN JOURNAL OF WOMENS STUDIES	235	256	91.8%
GENDER WORK AND ORGANIZATION	313	346	90.5%
SEXUALITIES	208	232	89.7%
GENDER AND EDUCATION	405	468	86.5%
NOUVELLES QUESTIONS FEMINISTES	116	141	82.3%
ASIAN JOURNAL OF WOMENS STUDIES	132	163	81.0%
WOMENS STUDIES INTERNATIONAL FORUM	484	604	80.1%
FEMINIST REVIEW	180	230	78.3%
FEMINIST THEORY	79	101	78.2%
FEMINISTISCHE STUDIEN	79	101	78.2%
INTERNATIONAL FEMINIST JOURNAL OF POLITICS	127	163	77.9%
POLITICS & GENDER	66	92	71.7%
FOCUS ON GENDER: PARENT AND CHILD CONTRIBUTIONS TO	5	7	71.4%
INDIAN JOURNAL OF GENDER STUDIES	102	147	69.4%
FEMINISM & PSYCHOLOGY	266	408	65.2%
GENDER MEDICINE	171	266	64.3%
HYPATIA-A JOURNAL OF FEMINIST PHILOSOPHY	110	177	62.1%
JOURNAL OF SEX RESEARCH	315	510	61.8%
SIGNS	241	504	47.8%
FEMINIST STUDIES	97	213	45.5%
AUSTRALIAN FEMINIST STUDIES	106	237	44.7%
FRONTIERS-A JOURNAL OF WOMEN STUDIES	52	179	29.1%
JOURNAL OF WOMENS HEALTH & GENDER-BASED MEDICINE	34	121	28.1%
PEDIATRIC GENDER ASSIGNMENT: A CRITICAL REAPPRAISAL	5	19	26.3%

Source: Compiled by Science-Metrix from WoS data (Thomson Reuters)

The precision (i.e. one minus the percentage of false positives, or irrelevant papers that were accidentally retrieved) of the dataset was examining the titles and abstracts of a random sample of 100 articles (Table 6). If the subject(s) of a paper did not relate to the above definition of GDRC, it was considered as a false positive. The precision was measured for the

dataset as a whole as well as for the fields in which GDRC papers are less likely to be found, such as in Information & Communication Technologies, Agriculture, Fisheries & Forestry, Engineering, and Earth & Environmental Science. The dataset as a whole has an excellent precision of 97%. The precision decreases a little for unusual fields, but since they are almost all above 70% and they do not represent a large proportion of the final dataset, this should not be a cause for concern.

Table 6. Precision of the GDRC dataset as a whole and for some fields in which relevant papers are less likely

Field	Precision	Sample size
Whole dataset	97%	100
Agriculture, Fisheries & Forestry	70%	50
Information & Communication Technologies	78%	50
Earth & Environmental Sciences	74%	50
Engineering	84%	50
Mathematics & Statistics	60%	50
Physics & Astronomy	80%	50

Source: Compiled by Science-Metrix from WoS data (Thomson Reuters)

All proportions are underestimated to a similar extent across countries for literature written in English. Care was taken not to bias the recall (i.e. the fraction of GDRC-relevant literature that was effectively retrieved and measured) in favour of specific countries.

Formula

Once the dataset is defined, the computation of the proportion of a country's publications integrating a gender dimension in its research content for a given year and field of science (FOS) is straightforward:

Table 7. Distribution of publications in the GDRC dataset across fields of science

Field	Papers	Share of field	Share of dataset
Total	212,673	N/A	100.0%
Clinical Medicine	91,162	2.9%	42.9%
Public Health & Health Services	33,546	10.5%	15.8%
Social Sciences	28,674	8.8%	13.5%
Psychology & Cognitive Sciences	20,961	7.9%	9.9%
Biomedical Research	10,685	0.9%	5.0%
Economics & Business	6,371	2.6%	3.0%
Communication & Textual Studies	5,363	5.2%	2.5%
Historical Studies	4,811	5.1%	2.3%
Philosophy & Theology	1,630	2.6%	0.8%
General Science & Technology	1,169	0.6%	0.5%
Information & Communication Technologies	1,108	0.2%	0.5%
General Arts, Humanities & Social Sciences	922	6.5%	0.4%
Enabling & Strategic Technologies	874	0.1%	0.4%
Visual & Performing Arts	715	3.4%	0.3%
Earth & Environmental Sciences	615	0.1%	0.3%
Engineering	537	0.1%	0.3%
Built Environment & Design	411	0.6%	0.2%
Biology	402	0.1%	0.2%
Chemistry	382	0.0%	0.2%
Mathematics & Statistics	267	0.1%	0.1%
Physics & Astronomy	234	0.0%	0.1%
Agriculture, Fisheries & Forestry	129	0.0%	0.1%
UNKNOWN	1,705	2.1%	0.8%

Source: Compiled by Science-Metrix from WoS data (Thomson Reuters)

As shown in Table 7, the fields that are present in high proportion relative to their size (column *Share of field*) are primarily from the Social Sciences and Humanities domain. Other than Clinical Medicine, the most represented fields (column *Share of dataset*) are Public Health & Health Services, Social Sciences, Psychology & Cognitive Sciences, and Biomedical Research, which is where one would expect most of the gender-related publications to be published. It is also interesting to note that the keyword query did catch a small number of articles from fields in which the presence of gender-related topics is less likely, such as Information & Communication Technologies, Earth & Environmental Sciences or Chemistry.

Table 8 shows the proportion of scientific publications that include a gender dimension in their content, by country and by field of science & technology over two time periods (2002–2005 and 2010–2013).

Table 8. Proportion of a country's scientific publications including a gender dimension in their research content, by field of science, 2002-2005 and 2010-2013

	Natural sciences		Engineering and technology		Medical sciences		Agricultural sciences		Social sciences		Humanities	
	02-05	10-13	02-05	10-13	02-05	10-13	02-05	10-13	02-05	10-13	02-05	10-13
World	0.1	0.2	0.0	0.1	2.8	3.9	0.0	0.0	6.8	7.2	3.9	3.9
EU-28	0.1	0.2	0.0	0.1	2.5	3.8	0.0	0.0	5.6	6.2	2.7	3.2
BE	0.2	0.2	0.0	0.1	2.4	3.5	0.0	0.0	4.3	5.1	1.6	2.4
BG	0.1	0.1	0.0	0.2	1.7	2.5	0.0	0.0	8.8	1.6	6.3	6.3
CZ	0.1	0.1	0.0	0.1	1.6	3.1	0.1	0.0	5.5	5.5	2.9	2.5
DK	0.3	0.3	0.1	0.1	4.5	5.3	0.0	0.0	2.6	4.6	2.5	3.6
DE	0.1	0.1	0.0	0.1	1.9	3.0	0.0	0.0	4.6	5.1	2.0	1.4
EE	0.3	0.1	0.0	0.3	3.8	7.3	0.0	0.0	8.1	6.8	0.0	1.7
IE	0.1	0.2	0.0	0.0	2.9	4.8	0.0	0.0	8.5	7.0	2.5	5.0
EL	0.1	0.3	0.0	0.1	2.6	4.3	0.0	0.0	3.8	4.9	4.4	3.0
ES	0.1	0.2	0.0	0.1	2.6	4.1	0.0	0.0	4.2	6.5	0.7	2.4
FR	0.1	0.1	0.0	0.0	1.9	2.8	0.0	0.0	3.2	4.1	1.6	1.9
HR	0.1	0.3	0.0	0.1	3.5	4.3	0.0	0.0	8.7	8.5	0.0	3.5
IT	0.1	0.2	0.0	0.1	2.3	3.1	0.0	0.0	3.5	4.4	1.4	1.6
CY	0.2	0.2	0.0	0.0	7.2	7.2	0.0	0.0	6.0	6.7	0.0	1.0
LV	0.1	0.3	0.0	0.0	1.2	2.5	0.0	0.0	16.7	3.5	0.0	7.7
LT	0.1	0.1	0.2	0.4	4.5	5.5	0.0	0.0	6.7	3.6	10.5	0.0
LU	0.4	0.0	0.0	0.0	4.2	5.2	0.0	0.0	3.6	5.4	0.0	2.9
HU	0.1	0.1	0.0	0.1	2.1	4.1	0.0	0.0	5.6	4.2	3.4	2.2
MT	2.2	1.3	0.0	0.0	4.4	7.1	0.0	0.0	6.7	3.7	0.0	0.0
NL	0.2	0.4	0.1	0.1	3.2	4.3	0.0	0.0	5.4	5.5	2.8	2.6
AT	0.1	0.2	0.1	0.2	2.1	3.4	0.0	0.0	4.4	6.6	4.2	1.8
PL	0.0	0.1	0.0	0.0	2.0	3.7	0.0	0.0	4.7	4.9	3.9	2.5
PT	0.1	0.2	0.0	0.1	2.3	5.6	0.0	0.0	3.8	4.7	1.6	2.0
RO	0.0	0.2	0.1	0.0	2.6	3.1	0.0	0.0	6.1	4.2	11.7	2.1
SI	0.0	0.2	0.0	0.1	2.8	4.6	0.4	0.0	7.0	7.1	5.2	2.1
SK	0.0	0.2	0.0	0.2	2.4	3.7	0.0	0.0	5.1	6.1	3.3	0.9
FI	0.2	0.4	0.0	0.1	4.5	6.0	0.0	0.0	8.3	8.5	1.2	4.0
SE	0.2	0.4	0.1	0.1	5.0	6.5	0.0	0.2	7.6	8.8	3.3	7.6
UK	0.2	0.3	0.1	0.1	2.6	3.8	0.0	0.1	6.6	7.0	3.9	4.9
IS	0.1	0.5	0.0	0.4	8.8	8.6	0.0	0.0	9.5	12.4	0.0	4.6
LI	0.0	0.0	0.0	0.0	0.0	9.2	z	z	0.0	0.0	z	z
NO	0.3	0.3	0.1	0.1	6.0	7.6	0.1	0.0	7.6	6.8	4.4	3.8
CH	0.1	0.2	0.0	0.1	2.4	3.5	0.0	0.0	5.8	4.8	2.5	2.7
ME	0.0	0.0	0.0	0.0	z	5.8	0.0	0.0	0.0	6.1	z	0.0
MK	0.0	0.4	0.0	0.0	2.1	4.4	0.0	0.0	0.0	5.4	33.3	16.7
AL	0.0	0.5	0.0	0.8	7.0	7.0	0.0	0.0	0.0	0.0	0.0	25.0
RS	0.0	0.1	0.0	0.1	0.0	5.0	0.0	0.0	z	4.3	z	3.1
TR	0.1	0.4	0.0	0.5	2.8	5.3	0.0	0.0	9.3	9.8	1.9	3.0
BA	0.0	0.3	0.0	0.0	7.9	7.6	0.0	0.0	11.5	5.3	z	7.7
FO	0.0	0.0	0.0	0.0	6.3	0.0	0.0	0.0	z	0.0	0.0	0.0
IL	0.1	0.2	0.0	0.1	3.4	4.6	0.0	0.0	9.6	8.0	2.2	2.6
MD	0.0	0.0	0.0	0.0	2.0	5.5	0.0	0.0	25.0	0.0	z	0.0

Notes: z = not applicable. All proportions are underestimated to a similar extent across countries for literature written in English. Care was taken not to bias the recall in favour of specific countries. This is because it is very difficult to extract 100% of the relevant literature using text-mining techniques without compromising accuracy. This is especially true for GDRC as the terminology used in the SSH is more generic than in other scientific areas.

Source: *She Figures 2015* (European Commission, 2016), originally computed by Science-Metrix using WoS data

A comparison between the figures worldwide and those in the EU-28 reveals that the propensity to include a gender dimension in research subject matter is similar to the world average within the EU-28 Member States. The breakdown by fields of science & technology shows that the gender dimension is most common in the social sciences in 2010–2013 (6% and 7% in the EU-28 and the world, respectively). The humanities and the medical sciences display a more modest share of publications with a gender dimension in 2010–2013 (3% to 4%), while the gender aspect is generally lacking or very minor in the agricultural sciences, engineering and technology, and natural sciences.

There is considerable country variation in the extent to which the gender dimension is addressed in national research outputs if the percentage of positive or negative departure from EU-28 or world level is considered; however, these departures (in percentage points) are generally small.

The overall small differences between time periods suggest that no major advances have been made in terms of addressing the gender dimension in research. The small differences nevertheless point to increases rather than decreases.

DISCUSSION

The propensity to integrate a gender dimension into research content is generally modest, but where changes have been observed over the two time periods investigated, these point towards an increase. Unsurprisingly, a gender dimension is most likely to be taken into account is in the social sciences, and least likely in the agricultural sciences, engineering and technology, and natural sciences. For further analysis of this indicator, readers are referred to the *She Figures 2015* publication (European Commission, 2016).

A database developed by Charité Berlin indexes sex- and gender-related literature in the biomedical field. Their methodology to identify relevant literature (Oertelt-Prigione, Parol, Krohn, et al., 2010) also makes use of keyword-based query terms followed by manual validation, however only 10 highly relevant keywords were used in their study. Despite differences in the exclusion and acceptance criteria, most of the terms were consistent with those used in the present study with the exception of sexual dimorphism and sexually dimorph*. Given the focus on animal studies, the inclusion of these terms was appropriate in their case as articles using these terms in their titles or abstracts tend to be about animals, whereas the present study focuses on humans. The exclusion of such terms in the current study, even if they retrieve a few relevant papers out of their total, is not critical given the high degree of redundancy in our approach; that is, the inclusion of a wide variety of search terms that often co-occur in publications. Indeed, if a keyword is omitted, a relevant publication including this keyword will likely be captured by one of the selected search expressions.

It is important to note that this is a newly developed indicator and that any reference or target about appropriate levels of the indicator is lacking. Although it is difficult to establish a target for what could be considered ‘adequate’ content, the observed shares generally appear low. This remains true even if the GDRC dataset captures only roughly 60% of the relevant literature, implying that there is room for further increases in the future. The results presented should hence be considered as baseline levels, allowing their evolution to be monitored in the

future. Additionally, relevant data at the project-level would help to further strengthen the ability to monitor changes in the inclusion of the GDRC over time, and could be linked more directly with policy initiatives in major funding programs such as Horizon 2020.

REFERENCES

- Archambault, É., Beauchesne, O. H., & Caruso, J. (2011). Towards a multilingual, comprehensive and open scientific journal ontology. In B. Noyons, P. Ngulube, & J. Leta (Eds.), *Proceedings of the 13th International Conference of the International Society for Scientometrics and Informetrics (ISSI)* (pp. 66–77). Durban, South Africa. Retrieved from <http://science-metrix.com/en/publications/scientific-publications#/en/publications/scientific-papers/towards-a-multilingual-comprehensive-and-open-scientific-journal>
- Campbell, D. (2015). *She Figures 2015: Comprehensive methodology—New research & innovation output indicators*. Brussels: European Commission.
- European Commission. (2016). *She Figures 2015*. Brussels. Retrieved from https://ec.europa.eu/research/swafs/pdf/pub_gender_equality/she_figures_2015-final.pdf#view=fit&pagemode=none
- European Commission, & Directorate-General for Research and Innovation. (2014). Gender equality in Horizon 2020.
- European Parliament and the Council of the European Union. (2013). Laying down the rules for participation and dissemination in “Horizon 2020 - the Framework Programme for Research and Innovation (2014-2020)” and repealing Regulation (EC) No 1906/2006. *Official Journal of the European Union*, 347.
- OECD. (2002). *Frascati Manual 2002*. Paris: Organisation for Economic Co-operation and Development. Retrieved from <http://www.oecd-ilibrary.org/content/book/9789264199040-en>
- OECD. (2007). Working party of national experts on science and technology indicators: Revised field of science and technology (FOS) classification in the Frascati Manual. Retrieved from <http://www.oecd.org/science/inno/38235147.pdf>
- Oertelt-Prigione, S., Parol, R., Krohn, S., Preißner, R., & Regitz-Zagrosek, V. (2010). Analysis of sex and gender-specific research reveals a common increase in publications and marked differences between disciplines. *BMC Medicine*, 8(1), 70. doi:10.1186/1741-7015-8-70



Gender Differences in Synchronous and Diachronous Self-citations

Gita Ghiasi^{*}, Vincent Larivière^{**}, Cassidy R. Sugimoto^{***}

^{*}g_ghia@encs.concordia.ca

Department of Mechanical and Industrial Engineering, Concordia University, 1455 De Maisonneuve Blvd. W.,
Montréal, QC, H3G 1M8, Canada

^{**}vincent.lariviere@umontreal.ca

École de bibliothéconomie et des sciences de l'information, Université de Montréal, C.P. 6128, Succ. Centre-
Ville, Montréal, QC, H3C 3J7, Canada

^{***}sugimoto@indiana.edu

School of Informatics and Computing, Indiana University, 1320 E. 10th St, Bloomington, IN, 47405, United
States of America

ABSTRACT

Citation rates are increasingly used as a currency of science, providing a basis to reward a scientist. Self-citations, an inevitable part of scholarly communication, may contribute to the inflation of citation counts and impose a considerable impact on research evaluation and academic career advancements. Self-citations are classified into two types in this study: synchronous self-citations (self-citations an author gives) and diachronous self-citations (self-citations an author receives). The main objective of this paper is to provide a comprehensive gendered analysis of synchronous and diachronous self-citations across all scientific disciplines. For this purpose, citation data of 12,725,171 articles published in 2008-2014 are extracted from Web of Science and are further scrutinized for articles of each gender. The findings reveal that men receive citations from their own papers at a higher rate than their women counterparts. They also tend to give more citations to their own publications. Gender gap in citation impact decreases when first-author's diachronous citations are eliminated in the impact analysis. However, the gap does not vary when all-authors' diachronous citations are excluded. The results of this research is important for effective gender-related policy-making in the science and technology arena.

INTRODUCTION

Women have been long susceptible to the “Matilda effect” in science (Rossiter, 1993)—the opposite of so-called “Matthew Effect”(Merton, 1968), an allusion to a well-known colloquialism “the rich get richer and the poor get poorer”—where the provisions of more credit to eminent scientists afflict women scientists by systematically repressing or ignoring their contribution to research and attributing their work to their male colleagues.

This means that the scientific discoveries of women get little to no credit for the same quality work as their male peers, only because of their gender. This consequence can expose a more prominent scientist—generally a male scientist—to ever-more resources and thereby more recognition and credit (for a similar work) than their contemporary female peers, leaving female scientists of the field unknown and invisible (Duran & Lopez, 2014).

As a result, women's contribution to science is often overlooked in the receipt of prestigious research award and grants (Lincoln, Pincus, Koster, & Leboy, 2012) and their research is less valued (Trix & Psenka, 2003; Wennerås & Wold, 1997). Women are commonly associated with

low-quality publications and are subject of lower collaboration interest (Knobloch-Westerwick, Glynn, & Hüge, 2013). Their publications are cited less frequently than their male peers after controlling for authorship positions (i.e. sole, first and last author) (Larivière, Ni, Gingras, Cronin, & Sugimoto, 2013), and affiliation, tenure status, methodology and context (Maliniak, Powers, & Walter, 2013).

Matilda effect has repeatedly shown to be present at citation level: although being published in journals with higher citation rates (Impact Factor), female-authored papers receive lower number of citations (Ghiasi, Larivière, & Sugimoto, 2015; Knobloch-Westerwick & Glynn, 2013; Larivière, 2014). Yet, citation rates are increasingly used as a currency of science—which mirror a base from which to reward a scientist (Merton, 1973)—and have become lamentably popular as the determinants of hiring, reappointment, tenure, promotion (Holden, Rosenberg, & Barker, 2005) and faculty salary (Toutkoushian, 1994), disfavoring women in their scientific research system.

Along these lines, self-citations may not only help inflate an author's citation counts, but impose a considerable impact on scholarly careers of academic researchers—one additional self-citation to a given paper attracts one extra citation from other researchers after one year and three extra citations after five years (Fowler & Aksnes, 2007). As citation rates are becoming a popular measure for research evaluations, self-citations are increasingly being used as a convenient means for manipulating the rewarding system of a scientist—paving the way for a researcher to gain more recognition and thus become more visible and influential.

Self-citations are the inevitable consequence of expanding on earlier study or furthering research in a specific field, but are also served as a tactic of manipulation for increasing a researcher's h-index, which is proved to have a positive impact on the academic ranking (Bartneck & Kokkermans, 2010) of a scientist. Hence, self-citations might add to the persistence of Matilda effect in science and play a major role in forming the scientific system that disfavors women in hiring and tenure procedures, salary decisions, and workplace advancements.

Gendered analysis of self-citations is very nascent and is limited to (Hutson, 2006; King, Correll, Jacquet, Bergstrom, & West, 2015; Susarla, Swanson, Lopez, Peacock, & Dodson, 2015). Among these studies, no significant effect between gender and self-citations have found (Hutson, 2006; Susarla et al., 2015). However, Hutson (2006) further scrutinized this finding and found that men cite themselves more often than women cite themselves (even when considering the rate of self-citations within the text of a paper). This is in line with the work of King et al. (2015) who also found an increasing gender gap in self-citations over the last 50 years.

Self-citations are generally categorized into two types (Aksnes, 2003), namely synchronous and diachronous self-citations. The former applies when the author cites his/her previous paper(s) in the paper that is being studied and the latter is when the paper that is being studied is cited by the author in one of his/her subsequent papers. Within the relevant literature, there is a gap in differentiating between these two types of self-citations and analyzing two statistics: cited by (or to) self and cited by (or to) all authors of the paper. This study tries to fill this gap and provides a comprehensive gendered analysis of synchronous and diachronous self-citations across all scientific fields.

METHOD

Article data is gathered from Thomson Reuters' Web of Science (WoS)—a comprehensive database of peer-reviewed publications and citations, which contains the Science Citation Index Expanded, the Social Science Citation Index and the Arts and Humanities Citation Index. The classification of scientific disciplines are based on that of the U.S. National Science Foundation's (NSF), which assigns each journal to only one discipline and one specific specialty and is necessary to avoid multiple counting of articles that are published in multidisciplinary journals. NSF classifies disciplines into 14 categories, namely Arts, Biology, Biomedical Research, Chemistry, Clinical Medicine, Earth and Space, Engineering and Technology, Health, Humanities, Mathematics, Physics, Professional Fields, Psychology, and Social Sciences.

A total of 12,725,171 articles are identified for the years 2008-2014 and are assigned to the aforementioned disciplines (Table 1). The focus of this study is on articles published after the year 2008, because WoS covers the full first names of authors from the year 2008 (i.e. essential for assigning gender to the authors).

Table 1. Number of articles per discipline over the period 2008-2014

Discipline	Number of Articles	Discipline	Number of Articles
Arts	142,015	Health	319,803
Biology	800,856	Humanities	574,585
Biomedical Research	1,365,778	Mathematics	324,398
Chemistry	1,059,575	Physics	885,975
Clinical Medicine	4,242,232	Professional Fields	348,092
Earth and Space	555,983	Psychology	262,391
Engineering and Technology	1,366,841	Social Sciences	476,647

Gender of WoS authors is further assigned by matching authors' given names with universal and country-specific existing name and gender databases, including U.S. Census, WikiName, Wikipedia, France and Quebec lists, and country-specific lists (which is explained in detail in (Larivière et al., 2013)).

Citation data is normalized for publication year and subject area, and is measured as the average yearly number of citations to a given paper from its year of publication to end of the year 2014, divided by the average yearly number of citations received by all papers published in the same year and in the same field.

Diachronous self-citations are identified by matching author names of citing articles with author names on a cited article (authors with the same last name and first name abbreviations). Citation data is then normalized, excluding diachronous self-citations received by both first author's and co-authors' papers. Finally, gender gap is calculated as the difference between average citation rates of articles first-authored by men and women relative to average citation rates of male first-authored articles.

Similarly, synchronous self-citations are identified where articles listed as references of a given paper have authors with similar name to authors of that paper (same last name and first name abbreviations), and are further grouped into two measures: citations to first-author's articles

and citations to articles written by any of the authors. Self-referencing rate is defined for each gender as the ratio of self-citations to total number of references and gender gap in self-referencing rate is then calculated. The smaller the percentage, the more equal researchers of each gender self-cite.

RESULTS

Citation impact of articles first-authored by male scientists is higher in all different disciplines (except Arts). Gender gap in citation impact is 9% for all scientific disciplines (Fig. 1) and is the highest in Biomedical Research, Psychology and Chemistry (14%). It is the lowest in Engineering and Technology and Biology (3%) and is non-existent in Arts (female-authored papers in arts receive more citations). However, the gender gap in citation impact slightly shrinks when excluding the first-author's diachronous self-citations, revealing more equal rate of citations received by articles of each gender. Interestingly, when eliminating all-authors' diachronous self-citations, the same rate for gender gap is reached as if all the self-citations are included. This applies to all scientific disciplines in all the years (Fig. 2) and might imply that male first-authored articles tend to receive citations from their subsequent publications at a higher rate than female first-authored articles. Nevertheless, when women are first authors, their papers might receive citations from their co-authors at a higher rate than that of their male counterparts—which, in part, explains the equal percentages for gender gap including and excluding all author's self-citations.

Figure 1: Gender gap in citation impact of scientific papers by discipline, including and excluding diachronous self-citations (2008-2014)

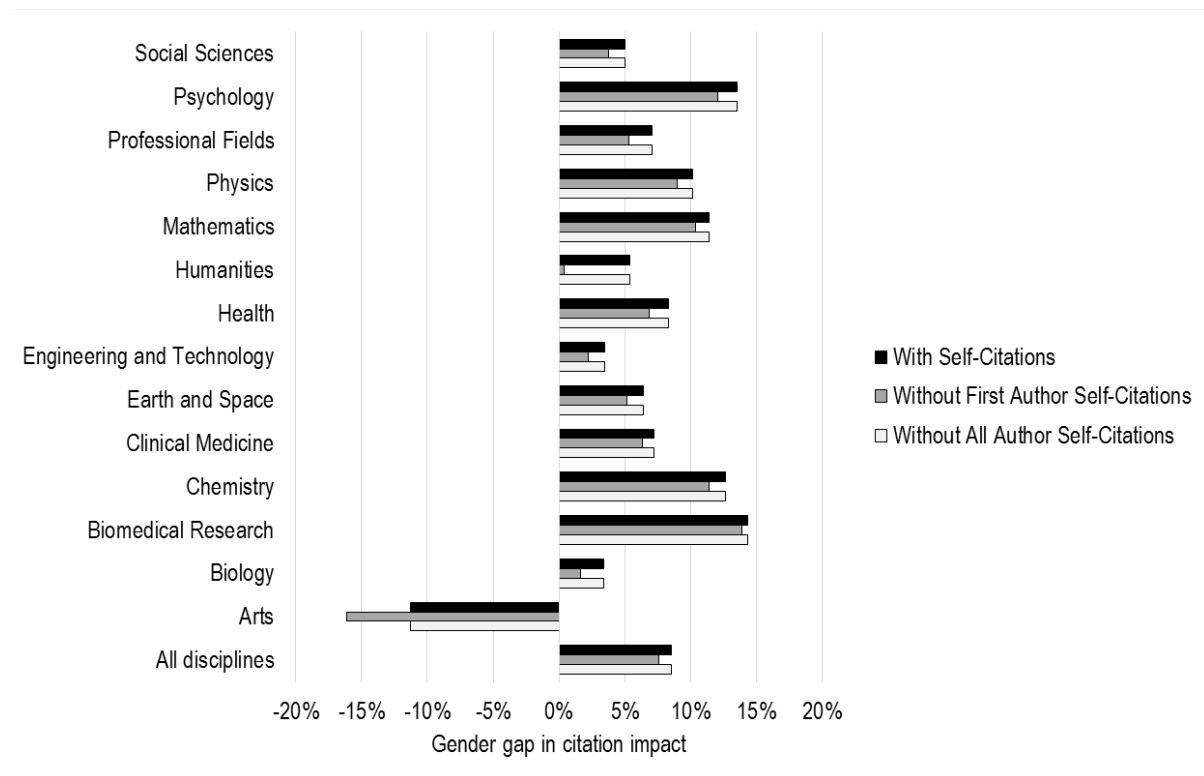
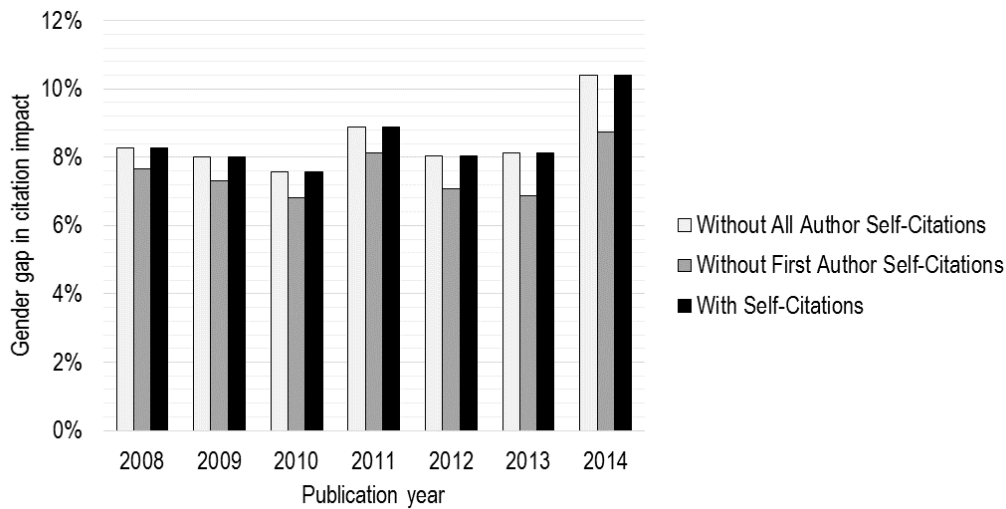


Figure 2: Gender gap in citation impact of scientific papers by year, including and excluding diachronous self-citations



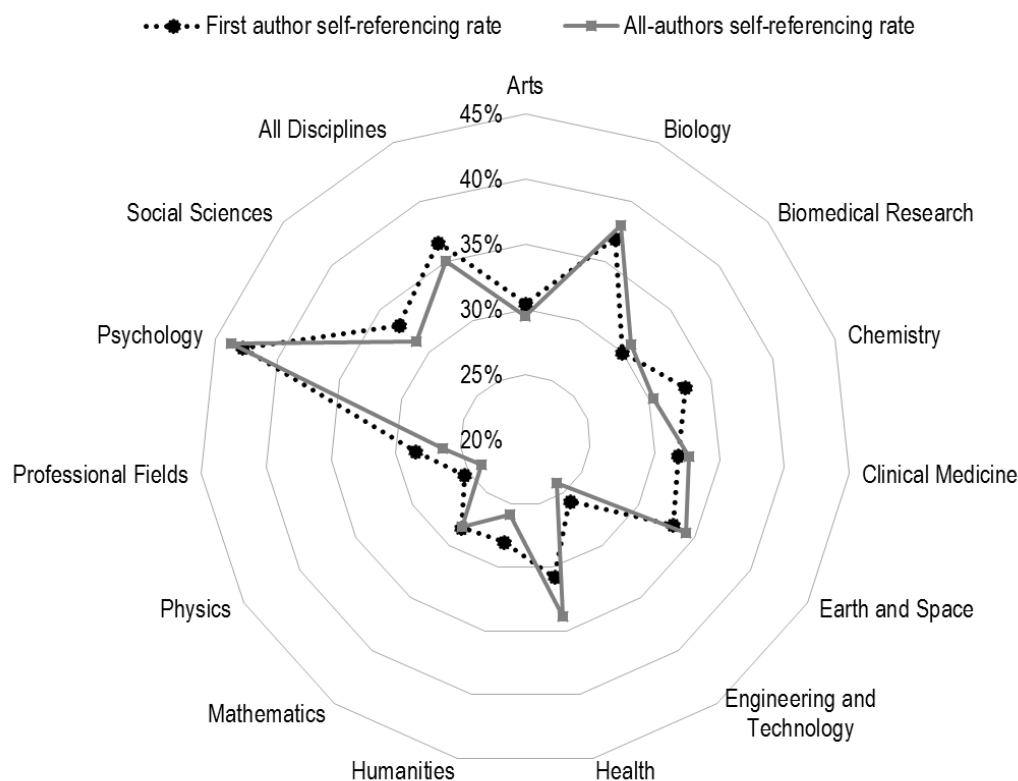
The analysis of synchronous self-citation rates (aka, self-referencing rates) shows that men, when listed as the first author of a paper, tend to cite their past works or any of the authors’ previous works at a higher rate than their female peers across all the disciplines (Table 2). The highest rate of self-referencing (for both female and male first-authored articles) is in Mathematics—which might be due to the fact that research in this particular field is expanding on earlier hypotheses and methods—and the lowest rate is in Biomedical Research and Clinical Medicine, showing that new and dynamic discoveries are furthering research in these areas.

Table 2. First author and all-authors synchronous self-citation (self-referencing) rate for female and male authored papers by discipline (2008-2014)

Discipline	First-author self-referencing rate		All-author self-referencing rate	
	Female	Male	Female	Male
Arts	3.3%	4.7%	4.6%	6.5%
Biology	2.1%	3.4%	3.5%	5.7%
Biomedical Research	1.5%	2.2%	3.1%	4.5%
Chemistry	2.3%	3.5%	4.0%	5.7%
Clinical Medicine	1.5%	2.3%	3.0%	4.5%
Earth and Space	2.5%	3.8%	4.3%	6.6%
Engineering and Technology	2.9%	3.8%	4.5%	5.9%
Health	2.2%	3.1%	3.4%	5.1%
Humanities	3.8%	5.2%	4.9%	6.6%
Mathematics	4.9%	6.9%	6.6%	9.2%
Physics	3.4%	4.5%	5.5%	7.3%
Psychology	2.3%	4.1%	3.6%	6.5%
Social Sciences	2.4%	3.6%	3.4%	5.0%
Professional Fields (Others)	1.8%	2.6%	2.7%	3.7%
All Disciplines	1.9%	3.1%	3.5%	5.3%

Gender gap in first-author self-referencing rate is 37% and in all-author self-referencing rate is 35% (Fig. 3). The lowest gender gap in self-referencing rate is in Physics and Engineering and largest is in Psychology. Surprisingly, psychology is among the fields with highest rate of female authorship (women represent ~50% of total authorship), whereas physics and engineering are the most male-dominated disciplines (women account for only 20% of total authorship) (Larivière, 2014). The lower gender gap in self-referencing rate might be associated with the selection effect in male-dominated fields—the fact that women need to be highly competent in order to survive or stay in in the most male-dominated fields—and their research, hence, serves as an important resource upon which new discoveries are grounded.

Figure 3: Gender gap in self-referencing rate by discipline (2008-2014)



DISCUSSION

Articles that are authored by a male author, tend to receive higher rate of citations from his subsequent publications. When first-author diachronous citations are removed from the citation impact analysis, the gender gap decreases. Nevertheless, the gender gap in citation impact does not vary when all-author diachronous citations are excluded. This shows that a paper that is authored by a women, might receive higher citation rates from her co-authors' publications than a paper authored by a man, which might be associated to gender differences in self-promotions: although women self-promote their own works at lower level and their publications receive lower recognition (citations) from the scientific community than those of their male peers, their work is promoted and recognized at higher rate by their immediate co-authors.

Narrowing the focus to synchronous self-citations or (self-references), it can be noted that men, in their papers, refer to their previous works at a higher rate than women scientists. However, the gender gap in self-referencing rate is the lowest in Physics and Engineering—the most male-

dominated disciplines—which is bound to the exceptional competence and expertise of women researchers in the most male dominated fields (Dryburgh, 1999; Ghiasi et al., 2015), where their work serve as a basis from which new studies are made.

The exposure of women scientists to Matilda effect shed light on the lower citation impact of their work (although being published in higher Impact Factor journals) (Larivière, 2014). One of the elucidations for gender differences in citations is that men receive citations from their own papers at a higher rate than their women counterparts and they tend to give more citations to their own publications. Citation patterns of an article have proved to fall under the Matthew effect in the sense that papers with high number of citations continue to be cited at a higher rate (Merton, 1968, 1988). Therefore, gender differences in self-citations play a major role in attracting more citations, have a direct impact on the h-index score of an author, and might contribute to gender inequality in evaluation, hiring, promotion and pay in academia.

The results of this research is thus of utmost importance for effective policy-making, redefining scientific reward and evaluation system with the use of gender equity measures. Recognition of women's contributions to scientific research can help identify mismatches in science and technology policies that can thwart gender parities in scientific performance, which subsequently gear toward a more equitable society.

REFERENCES

- Aksnes, D. W. (2003). Characteristics of highly cited papers. *Research Evaluation*, 12(3), 159–170.
- Bartneck, C., & Kokkermans, S. (2010). Detecting h-index manipulation through self-citation analysis. *Scientometrics*, 87(1), 85–98.
- Dryburgh, H. (1999). WORK HARD, PLAY HARD Women and Professionalization in Engineering—Adapting to the Culture. *Gender & Society*, 13(5), 664–682.
- Duran, A., & Lopez, D. (2014). Women from Diverse Backgrounds in the Science, Technology, Engineering, and Math (STEM) Professions: Retention and Career Development. *Impact of Diversity on Organization and Career Development*, 214.
- Fowler, J. H., & Aksnes, D. W. (2007). Does self-citation pay? *Scientometrics*, 72(3), 427–437.
- Ghiasi, G., Larivière, V., & Sugimoto, C. R. (2015). On the Compliance of Women Engineers with a Gendered Scientific System. *PloS one*, 10(12), e0145931.
- Holden, G., Rosenberg, G., & Barker, K. (2005). Bibliometrics: A potential decision making aid in hiring, reappointment, tenure and promotion decisions. *Social work in health care*, 41(3-4), 67–92.
- Hutson, S. R. (2006). Self-Citation in Archaeology: Age, Gender, Prestige, and the Self. *Journal of Archaeological Method and Theory*, 13(1), 1–18.
- King, M. M., Correll, S. J., Jacquet, J., Bergstrom, C. T., & West, J. D. (2015). *Men set their own cites high: Gender and self-citation across fields and over time*. Working paper. Retrieved from <http://www.eigenfactor.org/gender/self-citation/SelfCitation.pdf>.
- Knobloch-Westerwick, S., & Glynn, C. J. (2013). The Matilda Effect-Role Congruity Effects on Scholarly Communication: A Citation Analysis of Communication Research and Journal of Communication Articles. *Communication Research*, 40(1), 3–26.

- Knobloch-Westerwick, S., Glynn, C. J., & Huge, M. (2013). The Matilda Effect in Science Communication: An Experiment on Gender Bias in Publication Quality Perceptions and Collaboration Interest. *Science Communication*, 35(5), 603–625.
- Larivière. (2014, March). Femmes et sciences : les premières données mondiales valident l'inégalité | Acfas | magazine Découvrir | mars 2014. *Acfas.ca*. Retrieved October 26, 2014, from <http://www.acfas.ca/publications/decouvrir/2014/03/femmes-sciences-premieres-donnees-mondiales-valident-l-inegalite>
- Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature*, 504(7479), 211–213.
- Lincoln, A. E., Pincus, S., Koster, J. B., & Leboy, P. S. (2012). The Matilda Effect in science: Awards and prizes in the US, 1990s and 2000s. *Social Studies of Science*, 42(2), 307–320.
- Maliniak, D., Powers, R., & Walter, B. F. (2013). The gender citation gap in international relations. *International Organization*, 67(04), 889–922.
- Merton, R. K. (1968). The Matthew effect in science. *Science*, 159(3810), 56–63.
- Merton, R. K. (1973). The normative structure of science. *The sociology of science: Theoretical and empirical investigations*, 267.
- Merton, R. K. (1988). The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. *Isis*, 606–623.
- Rossiter, M. W. (1993). The Matthew Matilda effect in science. *Social studies of science*, 23(2), 325–341.
- Susarla, S. M., Swanson, E. W., Lopez, J., Peacock, Z. S., & Dodson, T. B. (2015). Does Self-Citation Influence Quantitative Measures of Research Productivity Among Academic Oral and Maxillofacial Surgeons? *Journal of oral and maxillofacial surgery: official journal of the American Association of Oral and Maxillofacial Surgeons*, 73(10), 1981.e1–7.
- Toutkoushian, R. K. (1994). Using Citations to Measure Sex Discrimination in Faculty Salaries. *The Review of Higher Education*, 18(1), 61–82.
- Trix, F., & Psenka, C. (2003). Exploring the color of glass: Letters of recommendation for female and male medical faculty. *Discourse & Society*, 14(2), 191–220.
- Wennerås, C., & Wold, A. (1997). Nepotism and sexism in peer-review. *Nature*, 387(6631), 341–343.



Mapping the author gender-distribution of disease-specific medical research

Jens Peter Andersen^{*}, Jesper Wiborg Schneider^{**} and Mathias Wullum Nielsen^{***}

*jpa@ps.au.dk

Danish Centre for Studies on Research and Research Policy, Aarhus University, Bartholins Allé 7, Aarhus, DK-8000 C (Denmark)

**jws@ps.au.dk

Danish Centre for Studies on Research and Research Policy, Aarhus University, Bartholins Allé 7, Aarhus, DK-8000 C (Denmark)

***mwn@stanford.edu

History Department, Stanford University, Stanford, CA-94305-2024 (USA)

ABSTRACT

This short paper responds to a recent call for attention to the “diversity challenge” in biomedical research, specifically with regard to gender diversity. The lack of diversity can be limiting for the progression of knowledge production, a viewpoint shared by both the European Commission, the League of European Research Universities and the National Institute of Health.

We study the gender distribution of authors in medical research, specifically mapped to disease-classifications using the Medical Subject Headings (MeSH) found in PubMed Medline. The dataset consists of 1,542,050 papers, spanning from 2008-2015, with full author gender and disease classification, with metadata from PubMed and author affiliation information from Web of Science. The combination of information on disease and gender allows us to map diversity issues to specific diseases. Our hypothesis is that one of the decisive factors for variations in gender distribution is the difference in diseases specific to the male and female bodies.

On the most general level, we find that the distribution of authors is highly skewed toward an overrepresentation of male researchers. The areas, which have the highest representation of female researchers, pertain to either the female body or diseases affecting the language or communication of patients. In contrast, medical specialty rather than the male body, specifically the topic of fractures (orthopedic surgery); define the areas with the highest degree of male authors per paper. With the overrepresentation of male authors, there is a risk of research on diseases more prevalent among men to become the norm, which can lead to detrimental effects on women’s health.

INTRODUCTION

In September 2015, two representatives from the American National Institutes of Health (NIH) published a perspective piece in *PNAS* calling for increased attention to the “diversity challenge” in biomedical research. Women and ethnic minority groups, so they argue, remain underrepresented in the higher echelons of medical research, and this “imbalance is limiting the promise of [the] biomedical enterprise for building knowledge and improving the nation’s health” (Valantine & Collins, 2015, p. 12240). Like many other science policy stakeholders, including the European Commission and the League of European Research Universities, the NIH subscribes to the idea that diversity enhances knowledge outcomes (European Commission, 2012; Maes, Gvozdanovic, Buitendijk, Hallberg, & Mantilleri, 2012). Yet, little is known about whether and to what extent this is the case for the medical sciences. NIH has therefore urged the research community to develop scientific approaches specifically designed to document the impact of diversity on the quality and outputs of academic medicine (Valantine & Collins, 2015).

With a specific focus on gender diversity, our study makes a first attempt to meet this request. More specifically, we develop a new approach for analysing topical variations in the gender distributions among authors of disease-related research papers in the biomedical literature, hereby opening a so far under-explored frontier in the scholarship on gender diversity and scientific knowledge outcomes. The scholarly focus on conventional performance measures, such as publication productivity, citation rates, and department rankings, provides some evidence linking gender diversity and scientific outcomes (see e.g. Campbell, Mehtani, Dozier, & Rinehart, 2013; De Saá-Pérez, Díaz-Díaz, Aguiar-Díaz, & Ballesteros-Rodríguez, 2015; Herring, 2013a, 2013b), but it has done little to elucidate deeper, qualitative knowledge issues. To address this gap and push the research agenda forward, we introduce a new approach for exploring to what extent and in what ways an increasingly gender diverse biomedical workforce may influence scientific knowledge outcomes by offering a broader variety of viewpoints and questions. Drawing on insights from political research, we hypothesize that the gender composition of author groups influences the topics addressed in biomedical research, with male scholars being more oriented towards diseases concerning the male body and women vice versa. The remainder of this article will be structured as follows: First, we discuss the theoretical conceptualization of gender diversity underpinning our approach and specify the hypothesis presented above. Second, we describe the data and elaborate the selected method for mapping the author gender-distribution of disease-specific medical research. Third, we account for the results, and fourth we conclude by discussing the main findings and reflecting on implications for policy.

Theoretical perspectives and conceptual clarifications

In accordance with the scholarship on cultural diversity management (Ely & Thomas, 2001; Merrill-Sands, Holvino, & Cumming, 2000), we see gender as an identity category shaping cognition, experience and perspective. Gender diversity, following this approach, represents the conglomeration of worldviews, orientations and interests that women, men and other genders bring into the scientific enterprise. By adopting this conceptualization, we are not arguing that gender categories should be comprehended along essentialist lines. Yet, inspired by Bourdieu’s habitus concept (Bourdieu & Wacquant, 1992, p. 138), we do see gendered aspects of identity as embodied ‘ways of being’ acquired through socialization (McLeod, 2005); and human beings

carry these ‘ways of being’ into scientific organizations with potential implications for knowledge production. This theoretical assertion finds some support in the social science literature. Studies in political science remind us that not only the compositional effects of gender diversity (GD), but also women’s contributions in particular matter for organizational/societal outcomes. Schwindt-Bayer & Mishler (2008), for instance, use data from 31 democracies and find women’s descriptive representation among legislators to positively influence parliamentary responsiveness to policy concerns such as maternity leave, social and political equality and marital equality in law (see also Bratton & Ray, 2002; Childs, 2005; Lovenduski & Norris, 2003). What we can derive from this branch of scholarship is that representation matter in decision-making; not only symbolically (i.e. as a symbol of democratic inclusiveness) but also substantially (i.e. the passing of ‘women friendly’ policies).

We hypothesize that similar patterns moderated by gender can be detected in the biomedical sciences. A hypothetical example relates to sexual dysfunction diseases: Women biomedical researchers may be overrepresented among authors of scholarly papers addressing female sexual dysfunction diseases, whereas male dominated author groups could be expected to be more oriented towards male sexual dysfunction diseases.

We will pursue these conjectures more closely in future studies; but to do that we need to construct a useful data set, where we map the author gender-distribution of disease-specific medical research. The remainder of this brief methodological proceeding will outline and discuss this initial mapping attempt.

MATERIALS & METHODS

Data collection and MeSH classification

Data for this study are gathered from the Medline database using PubMed, which is one of the largest bibliographical databases indexing medical research. The use of PubMed Medline has two distinct advantages over other databases in the current context: data are freely available to download and all papers, which are indexed in the Medline-section of PubMed, are assigned a number of MeSH subject headings.

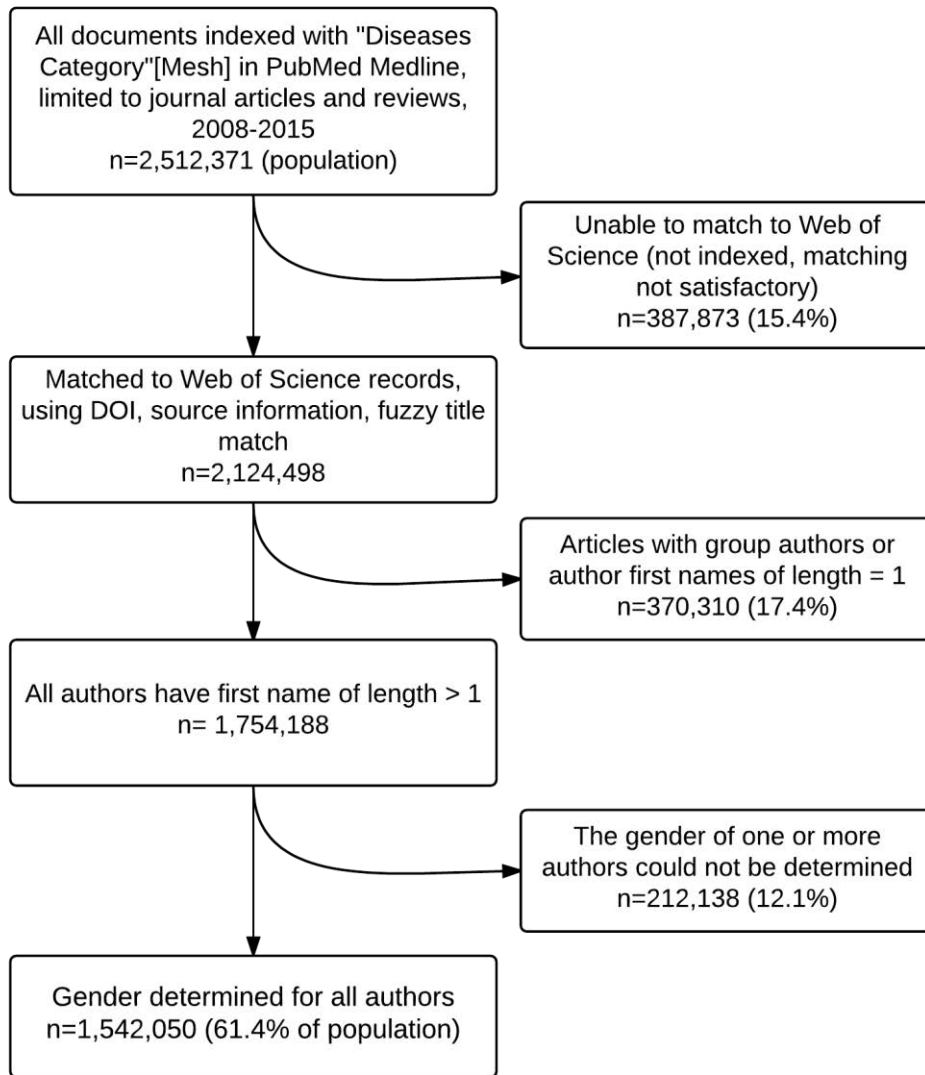
The MeSH thesaurus is a professional, hierarchical, controlled vocabulary used to index Medline records at the National Library of Medicine (NLM). There is a parent-child relationship so that all MeSH-terms must have at least one parent, but can have more than one, i.e. each child node may have more than one entry in the hierarchy, even at different levels. The MeSH thesaurus is also freely accessible in a machine-readable format using the Entrez e-utils¹. This option was used for this study in order to obtain all MeSH-terms subordinate to the “Diseases Category” term, including their links. Correspondingly, article metadata were gathered for those articles indexed with any of these terms, which is possible in PubMed by simply querying for “Diseases Category”[MeSH], as PubMed will automatically “explode” this query. We limited our search to papers published after January 1st 2008, up to December 31st 2015, resulting in 2,512,371 papers. Of these papers, 2,124,498 (84.5%) were matched to records indexed in Web of Science (WoS). For this purpose, we used the modified WoS database maintained by the CWTS at Leiden University. All WoS records for the same time period were compared against the PubMed set, using first DOI, then ISSN, journal names, pagination, volume and fuzzy title matches using relative Levenshtein distance. While it is possible to search PubMed accession numbers (PMID) in the WoS interface, this was not

¹<http://www.ncbi.nlm.nih.gov/books/NBK25501/>

This work is licensed under a Creative Commons License: Attribution-NonCommercial-NoDerivatives 4.0 International.

deemed a viable approach due to query limitations. Further, tests indicated that 2.5% additional records at most could be expected found using this approach. An overview of the data exclusion related to this process and to the subsequent assignment of gender to author names, as detailed in the next section, is provided in Figure 1.

Figure 1 - Flowchart of data in- and exclusion



Determining the gender of authors

To determine the gender of scholarly authors, we used the Gender-API (2016). Gender API draws on data from social media websites to assign gender specifications to first names, while accounting for variations across countries (e.g. Andrea is, for instance, a typical male name in

Italy but a female name in England).² While there may be more accurate connections between an author's first name and country of origin than her/his country of institutional affiliation, only the latter is available to us.

Information on first names and country of institutional affiliation was retrieved using the CWTS entry to WoS. While WoS has included information on first names since 2007, this information is not provided for all papers, and 370,310 papers were excluded from our data set for this reason. Further, 212,138 papers had at least one author for which it was not possible to determine the gender, leaving a total of 1,542,050 papers for the period 2008-2015, indexed with a subterm for the MeSH "Diseases Category", matched to WoS, for which the gender of all authors was determined.

Gender indicator

From Gender-API, we are provided with information on the accuracy of the gender classification of any name and country pair, which we convert into a probability of a name belonging to a female researcher, denoted f . We use this probability to calculate a weighted indicator, f_w , which is the mean value of f for all authors of a paper. These values will range between 0 and 1, with values closer to 1 indicating a higher proportion of female authors. When using this indicator, we operate purely on aggregate levels. This is important due to the insecurity related to the classifications as registered in the accuracy measurement. A paper with $f_w = 0.8$ could, for instance, be authored by two female researchers, while another all-male paper might have $f_w = 0.2$.

These article-level scores can then be aggregated for each MeSH-keyword as average scores. We have chosen this approach, rather than dividing the total number of female authors by the total number of authors (mean-of-sums versus sum-of-means) as we want to express the ratio of female authors per paper and not the overall ratio. Both approaches are however meaningful in each their way.

We evaluate the hypothesis by mapping the mean f_w scores for MeSH-terms on the most general level and hereafter provide specific examples for selected subareas. These subareas have been selected based on their gender composition and topic specifications (diseases related to the male and female body).

RESULTS

A total of 4,709 MeSH terms were gathered subordinately to the diseases category. The mean f_w for all these is 0.35, showing a clear bias towards male-dominated author groups, with a distribution very close to Gaussian (not shown). The terms are displayed as a graph network in Figure 2, with nodes representing MeSH-terms, and vertices the hierarchical structure. The network visualizes the complexity of the hierarchy and the clear overrepresentation of male-dominated research areas.

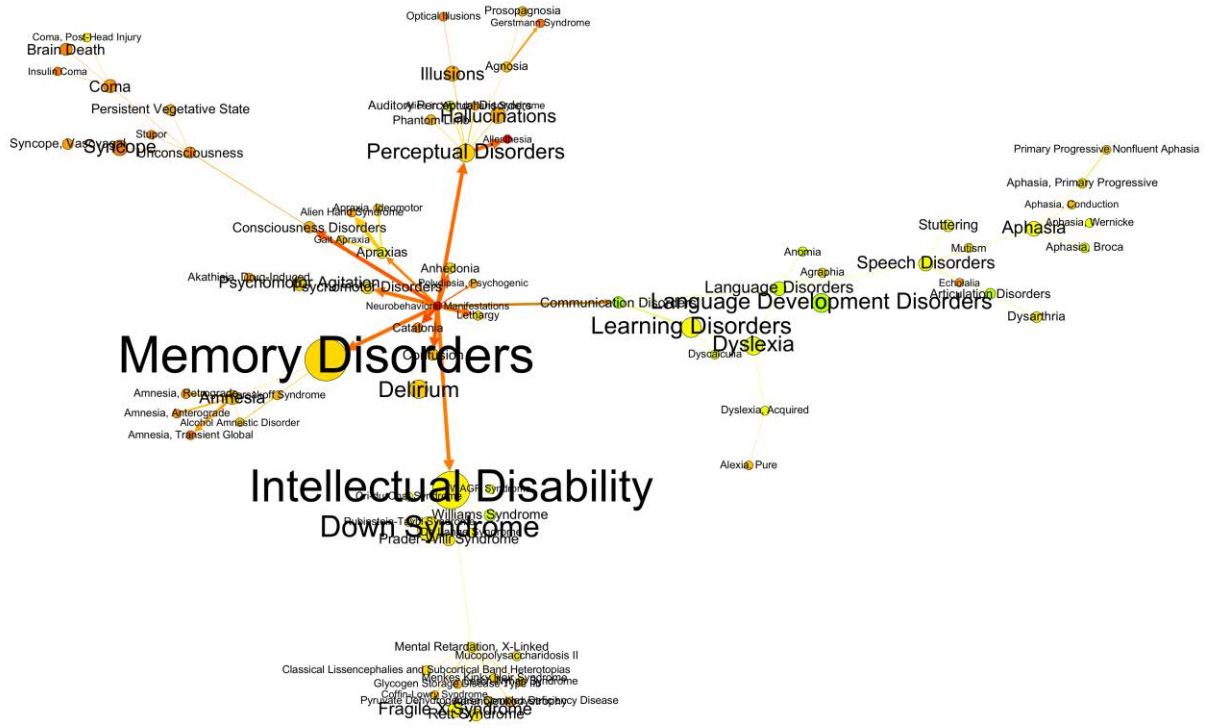
² The Gender API is a commercial tool, and the underlying process is unknown to us. Yet, random test samples from our data confirm its accuracy. Optimally, in the future this approach will be combined with the thesaurus-based method of Larivière, Ni, Gingras, Cronin, & Sugimoto (2013).

Figure 2 – Map of all MeSH terms colored by aggregated f_w values, using a color gradient starting from red ($f_w = 0$), to yellow ($f_w = 0.5$) and green ($f_w = 1$). Yifan Hu layout is used to position MeSH terms relative to each other.



Figure 3 depicts a selected subarea from the total map (Figure 2) related to the MeSH-term “Neurobehavioral manifestations” and its subordinate terms. This subarea includes MeSH-terms with both male and female dominance. Further, it illustrates clear, gender-based variations in preferences for research areas even in closely related topics.

Figure 3 - Map of MeSH terms subordinate to “Neurobehavioral manifestations”, colored by aggregated f_w values, using a color gradient starting from red ($f_w = 0$), to yellow ($f_w = 0.5$) and green ($f_w = 1$). Yifan Hu layout is used to position MeSH terms relative to each other.



In Figure 4, we show MeSH terms subordinate to female urogenital diseases, while Figure 5 depicts the corresponding male diseases. There is a large overlap between the maps. Especially kidney diseases and some sexually transmitted diseases are present in both areas. It is, however, obvious that the gender specific diseases (e.g. those pertaining to pregnancy or prostatic cancer) have different colors, meaning that the female diseases are relatively more likely to be investigated by female researchers and vice versa.

Figure 4 - Map of MeSH terms subordinate to “Female urogenital diseases”.

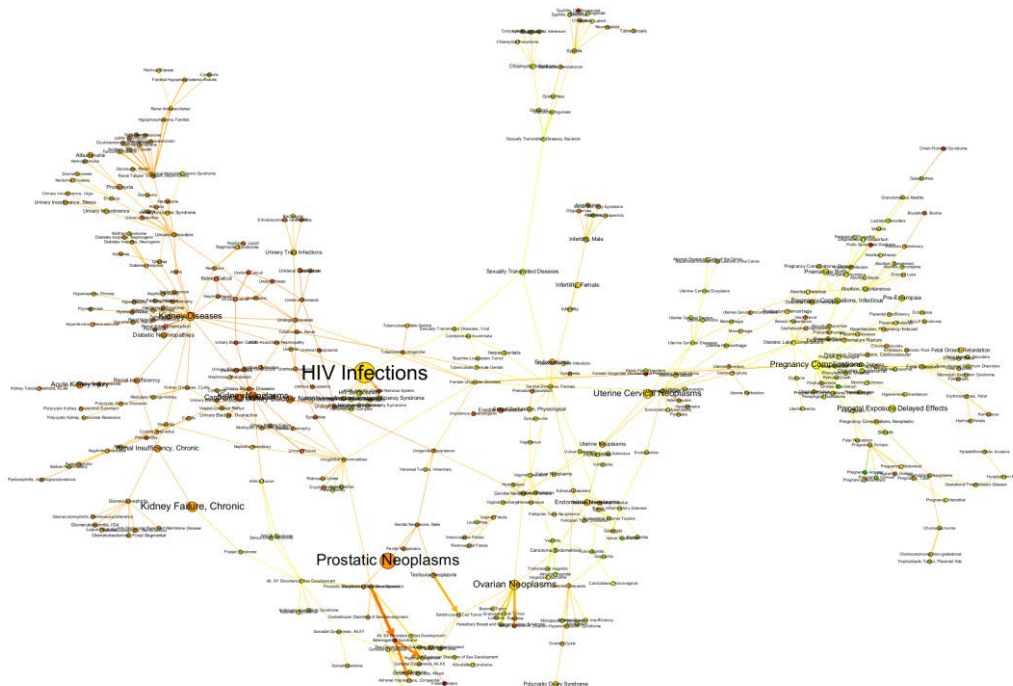
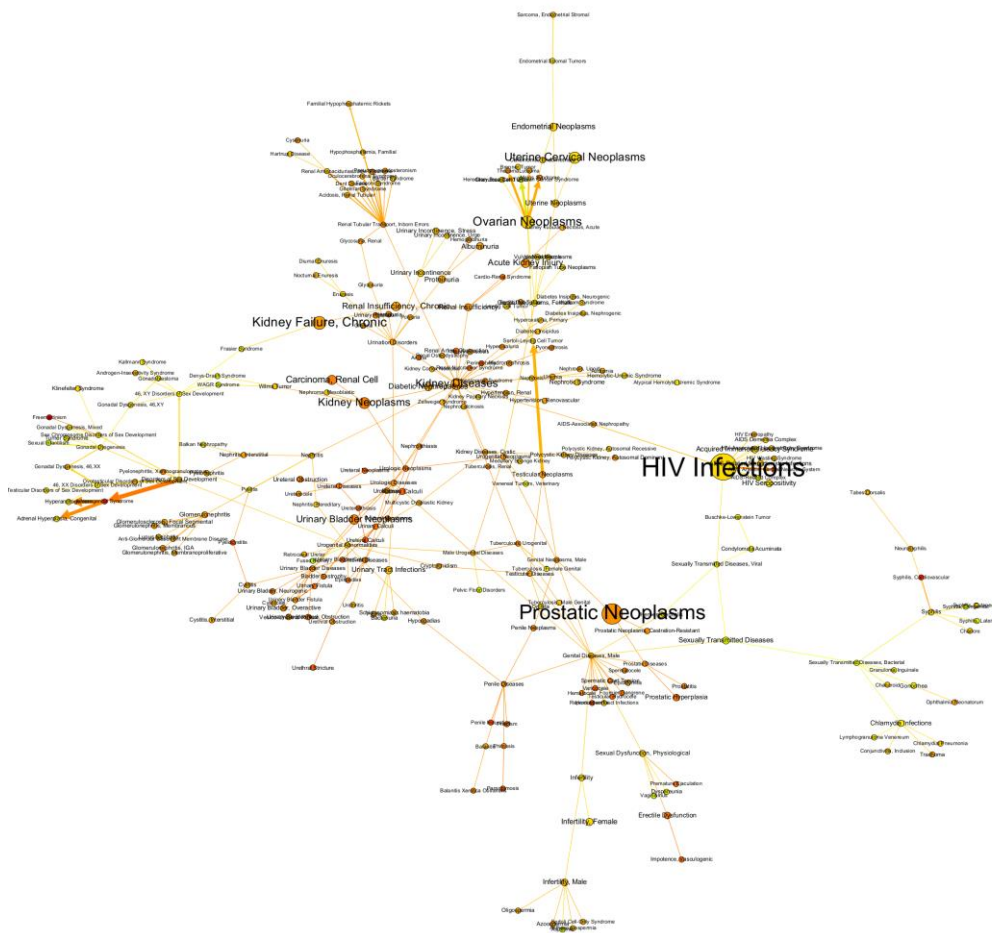


Figure 5 - Map of MeSH terms subordinate to “Male urogenital diseases”.



Having established these differences, we see a very clear orientation towards diseases regarding the female body among female dominated author groups. In Table 1, we show the 30 MeSH terms with the highest and lowest f_w value (and at least 50 uses in the dataset). A considerable amount of the female-dominated MeSH-terms are related to diseases specific to the female body, while the remainder relates to communication, language and speech related disorders. For the male topics, bone fractures stand out as the most prevalent, while an orientation towards diseases specific to the male body appears less prevalent.

Table 1 – Top-30 MeSH terms with at least 50 keyword-assignments for “Female” and “Male” categories. The categories are determined as having the highest and lowest f scores respectively.

“Female”			“Male”		
MeSH term	n	f_w	MeSH term	n	f_w
Communication Disorders	380	0.66	Periprosthetic Fractures	265	0.12
Depression, Postpartum	1163	0.66	Knee Dislocation	114	0.13
Language Development Disorders	1374	0.65	Intra-Articular Fractures	258	0.13
Vulvar Lichen Sclerosus	73	0.64	Fractures, Ununited	619	0.13
Articulation Disorders	265	0.63	Spondylolysis	95	0.13
Hereditary Breast and Ovarian Cancer Syndrome	51	0.62	Endoleak	374	0.13
Labor Pain	171	0.61	Hallux Rigidus	93	0.14
Bulimia	539	0.6	Pancreatic Fistula	456	0.14
Vulvodynia	115	0.58	Spondylolisthesis	529	0.14
Pediatric Obesity	899	0.58	Iliac Aneurysm	237	0.15
Williams Syndrome	391	0.58	Decompression Sickness	228	0.15
Anomia	105	0.58	Spinal Osteophytosis	81	0.15
Language Disorders	712	0.57	Fractures, Open	382	0.15
Vulvar Diseases	244	0.57	Humeral Fractures	641	0.15
Abortion, Incomplete	55	0.57	Bone Malalignment	510	0.15
Premenstrual Syndrome	311	0.56	Fractures, Comminuted	415	0.15
Trichomonas Vaginitis	171	0.56	Fractures, Malunited	323	0.15
Pelvic Floor Disorders	145	0.56	Osteochondritis Dissecans	230	0.15
Aphasia, Broca	146	0.55	Shoulder Fractures	490	0.15
Fetal Alcohol Spectrum Disorders	531	0.54	Ossification of Posterior Longitudinal Ligament	144	0.15
Menopause, Premature	118	0.54	Epiphyses, Slipped	81	0.15
Hot Flashes	657	0.54	Endolymphatic Hydrops	75	0.15
Sexually Transmitted Diseases	2359	0.54	Posterior Tibial Tendon Dysfunction	74	0.15
Vaginitis	90	0.53	Ureteral Calculi	527	0.15
Dyscalculia	58	0.53	Ulna Fractures	286	0.15
Dyslexia	1279	0.53	Shoulder Dislocation	556	0.16
Sexually Transmitted Diseases, Bacterial	141	0.53	Failed Back Surgery Syndrome	81	0.16
Auditory Perceptual Disorders	253	0.53	Joint Deformities, Acquired	199	0.16
Dyslexia, Acquired	58	0.53	Femoracetabular Impingement	423	0.16
Speech Disorders	779	0.53	Hammer Toe Syndrome	50	0.16

DISCUSSION

To briefly summarize, our preliminary analysis partly confirms our hypothesis. We show a clear relationship between the gender composition of author groups and the disease-related topics addressed in the biomedical research literature. Moreover, in line with our predictions, female dominated papers appear more likely to be oriented towards diseases specific to the female body. The research areas with the highest proportion of male authors, however, appear less “gendered” than expected. In most cases, the male-dominated areas pertain to the field of orthopaedic surgery; a medical speciality with around 90% male residents in the US (Nguyen, Amin, Vail, Pietrobon, & Shah, 2010). This finding highlights the importance of dedicating further attention to variations within and across medical specialties in a more detailed and elaborated version of the paper.

Obviously, the association between gender-author composition and topic selection is moderated by numerous factors left unnoticed in this paper. Existing literature, for instance, highlight considerable national differences in the gender composition of biomedical research staff (see e.g. European Commission, 2015); and since certain disease-related topics may be more prevalent in some countries than others, this should be taken into account in the subsequent steps of the analysis. Moreover, some research areas in the biomedical literature are considered more prestigious than others, which may play an important part in this regard. Following the sociological literature on gender and labour market stratification (Reskin & Roos, 2009), one may expect women to be ‘ghettoized’ into lower-status areas of research. One way of measuring this could be to compare the amounts of funding devoted to male and female dominated disease topics. Another way could be to look at the average number of citations accrued by papers published in male and female dominated research areas. Finally, it seems crucial to account for the gender of senior authors, since senior authors usually take the lead in identifying, planning and developing the research topics and questions.

All in all, our study makes a clear case for the relevance and importance of devoting further attention to author-based gender differences in biomedical research. As illustrated by our preliminary analysis, gender diversity matters for knowledge production; and the general implications of maintaining a global research system, where the vast majority of biomedical research is still conducted by men, may be more detrimental to women’s health than previously suspected.

REFERENCES

- Bourdieu, P., & Wacquant, L. J. (1992). *An invitation to reflexive sociology*. University of Chicago press.
- Bratton, K. A., & Ray, L. P. (2002). Descriptive representation, policy outcomes, and municipal day-care coverage in Norway. *American Journal of Political Science*, 46(2), 428–437.
- Campbell, L. G., Mehtani, S., Dozier, M. E., & Rinehart, J. (2013). Gender-heterogeneous working groups produce higher quality science. *PloS One*, 8(10), e79147. <http://doi.org/10.1371/journal.pone.0079147>
- Childs, S. (2005). The Substantive Representation of Women: The Case of the Reduction of VAT on Sanitary Products. *Parliamentary Affairs*, 59(1), 10–23. <http://doi.org/10.1093/pa/gsj003>

- De Saá-Pérez, P., Díaz-Díaz, N. L., Aguiar-Díaz, I., & Ballesteros-Rodríguez, J. L. (2015). How diversity contributes to academic research teams performance. *R&D Management*, ahead of print. <http://doi.org/10.1111/radm.12139>
- Ely, R. J., & Thomas, D. A. (2001). Cultural diversity at work: The effects of diversity perspectives on work group processes and outcomes. *Administrative Science Quarterly*, 46(2), 229–273.
- European Commission. (2012). Communication from the commission to the European Parliament, the Council and The European Economic and Social Committee and the Committee of the Regions: A reinforced European Research Area Partnership for Excellence and Growth.
- European Commission. (2015). She Figures 2015 - Leaflet. Retrieved March 8, 2016, from https://ec.europa.eu/research/swafs/pdf/pub_gender_equality/she_figures_2015-leaflet-web.pdf
- Gender API. (2016). Gender API - Determines the gender of a first name. Retrieved March 8, 2016, from <https://gender-api.com/>
- Herring, C. (2013a). Critical diversity and departmental rankings in chemistry. *Abstracts of Papers of the American Chemical Society*, 246.
- Herring, C. (2013b). Critical Diversity and Departmental Rankings in Sociology. *The American Sociologist*, 44(3), 267–281. <http://doi.org/10.1007/s12108-013-9183-8>
- Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature*, 504(7479), 211–213. <http://doi.org/10.1038/504211a>
- Lovenduski, J., & Norris, P. (2003). Westminster Women: The Politics of Presence. *Political Studies*, 51(1), 84–102. <http://doi.org/10.1111/1467-9248.00414>
- Maes, K., Gvozdanovic, J., Buitendijk, S., Hallberg, I. R., & Mantilleri, B. (2012). *Women, research and universities: excellence without gender bias*. Leuven, Belgium.
- McLeod, J. (2005). Feminists re-reading Bourdieu: Old debates and new questions about gender habitus and gender change. *Theory and Research in Education*, 3(1), 11–30. <http://doi.org/10.1177/1477878505049832>
- Merill-Sands, D., Holvino, E., & Cumming, J. (2000). *Working with diversity: A framework for action*.
- Nguyen, L., Amin, N. H., Vail, T. P., Pietrobon, R., & Shah, A. (2010). Editorial: A Paucity of Women Among Residents, Faculty, and Chairpersons in Orthopaedic Surgery. *Clinical Orthopaedics and Related Research*, 468(7), 1746–1748. <http://doi.org/10.1007/s11999-009-1125-y>
- Reskin, B. F., & Roos, P. A. (2009). *Job queues, gender queues: Explaining women's inroads into male occupations*. Temple University Press.
- Schwindt-Bayer, L. A., & Mishler, W. (2008). An Integrated Model of Women's Representation. *The Journal of Politics*, 67(02). <http://doi.org/10.1111/j.1468-2508.2005.00323.x>
- Valantine, H. A., & Collins, F. S. (2015). National Institutes of Health addresses the science of diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 112(40), 12240–2. <http://doi.org/10.1073/pnas.1515612112>



Indicators for constructing scientific excellence: ‘Independence’ in the ERC Starting Grant¹

Helene Schiffbaenker* and Florian Holzinger**

*helene.schiffbaenker@joanneum.at

POLICIES, JOANNEUM RESEARCH, Sensengasse 1, 1090 Vienna, Austria

**Florian.Holzinger@joanneum.at

POLICIES, JOANNEUM RESEARCH, Sensengasse 1, 1090 Vienna, Austria

SUMMARY

Scientific excellence is of increasing relevance for assessing and funding research. Yet the definition of excellence is unclear and excellence difficult to measure. Classical indicators like scientific outcome and impact have been identified to have various limitations, in particular from a gender perspective. So other indicators are applied, like independence. In this paper we discuss various aspects of applying independence as indicator for excellence in ERC peer review panels. In core we analyse how independence is formally defined and how it is applied in practice, with focus on its gendered effects. Further, potential and limitations of independence as criterion for excellence are discussed.

INTRODUCTION

The construction of scientific excellence has been widely discussed in research, as it is of increasing relevance for selecting the best quality of science, specifically relevant in research funding. While excellence is related to the meritocratic understanding of science that success is based on individual performance and merit only, recent research has demonstrated that excellence is socially constructed (O’Connor and O’Hagan 2015, Rees 2011, Lamont 2009, Brouns and Addis 2004), depending on scientific disciplines, cultural context, individual preferences and gender stereotypes (Heilman et al. 2015). The ideal academic is often constructed “as a lone, **independent** individual, who is self-protective, competitive, ruthless and not that collegiate or supportive of colleagues and students.” (Bleijenbergh/Engen/Vinkenburgh 2012: 24).

Gender research has revealed that indicators typically used to describe excellence are gendered (Rees 2011, van den Brink and Benschop 2012). Performance indicators such as publications have widely been criticised for not being gender neutral, since the lower amount of time women are able to spend due to care responsibilities and unpaid work lowers their productivity (van den Brink and Benschop 2012). Aksnes et al. (2011) have shown that the

¹ This work was supported by Peter van den Besselaar, Claartje Vinkenburgh and Lucia Polo, many thanks for this supportive cooperation.

hierarchical position explains bias more than gender, but as women are in general lower positioned in the science field they again publish less.

In order to improve the assessment of scientific excellence, ‘independence’ was introduced as an alternative indicator, covering social, topical and geographical independence (van den Besselaar et al. 2012).

The European Research Council (ERC) funds frontier research, which should have ground breaking impact on science. But over all the years and grants, success rates have been lower for female applicants. In a study of the ERC peer review process, we have analysed how criteria for assessing excellence are defined and put into practice by ERC panel members and if there is a gender bias.

In this paper, we focus on one specific aspect of excellence, which is independence. We first discuss if and how independence is defined by the ERC in its official documents. Next, we analyse how the formal understanding of independence is applied in practice, also discussing the link between formalisation and practice of excellence. Next, general practices when deploying ‘independence’ as indicator for excellence are presented. We analyse the discourse on independence by ERC panel members, focussing on expectations, potential and limitations. Here, the gender dimension becomes relevant as we describe the different attribution of independence to female and male applicants. Finally, we discuss if independence as a criterion itself is gender biased.

DATA

This paper is based mainly on qualitative interviews conducted with 32 members of ERC peer review panels between March and June 2015. We used the ERC Starting Grant (StG) 2014 as a case to study. All interviews have been recorded, transcribed and analysed using the software package MAXQDA. Additionally, ERC policy documents have been analysed to find out how independence is formalised. Furthermore was an online survey sent to all applicants who had given informed consent (n= 3.030) to be contacted to capture their perspective. Another source of our analysis were evaluation reports which are written assessments for each applicant produced by ERC panel members. This allowed us to investigate if assessment and feedback words are used differently for female and male applicants.

FINDINGS

Independence at the formal level

First, we look at the formal understanding of independence specified in the various ERC policy documents. This formalisation of independence covers the criterion and potential indicators for measuring the criterion.

The work programme 2014 refers to independence when describing the objectives of the StG: *“ERC Starting Grants (...) are designed to support excellent Principal Investigators (PIs) at the career stage at which they are **starting** (...) their own independent research team or programme”* (EC 2013: 20).

This illustrates that the StG should be used by a PI to start an own independent topic – which is not fully in line with another definition in the work programme (EC 2013: 20) “*A competitive **Starting Grant** candidate **must have already shown the potential** for research independence and evidence of maturity. For example, it is expected that applicants will have produced **at least one important publication without the participation of their PhD supervisor**”.*

Here, “at least one publication without the PhD supervisor” is seen as evidence of maturity, as expected advancement in a researcher’s scientific development and can be perceived as an indicator to measure the potential for independence. No further indicators are mentioned in any documents. So the formal understanding of independence refers to its social dimension as it asks the PI to demonstrate the ability to publish without the (former) supervisor. It should be stressed that other aspects of independence, like geographic independence (= mobility), are not formalised as a criterion and should not be considered in the ERC grant selection process.

Based on these formal definitions, all panel members and external reviewers are supposed to discuss independence when assessing excellence in both stages of the peer review process. In the remote phase each reviewer has to assess the PI’s *intellectual capacity and creativity*. Related to independence the reviewers are asked: “*To what extent does the PI provide evidence of creative independent thinking?*” (EC 2014: 2).

2 Independence in practice

Analysing the application of independence by panel members made evident that while only one indicator was formally proposed in ERC policy documents, panel members mentioned using several further indicators for assessing independence:

Independence is the composition of the team, and what is the latest publication – whether the applicant was the last author in the last publication is the direct indication of independence. And whether his team is strong enough, if he’s sharing PhD students or Postdocs with someone else or whether he has his own. Also funding; if you don’t have funding you can’t be independent. If you depend on someone else’s money, forget about your independence. (Panel member 21, LS, female)

After having listed all these different indicators, the interviewee points to what in fact is perceived as the core of independence, this is the ‘topical independence’.

There are very clear and hard criteria. I can tell you immediately – within two minutes – whether a person is independent or not. ... I would say [it would be] the research line or area compared to what this person’s supervisor for Postdoc or PhD has been doing before. That’s the major criteria when it comes to independence. (Panel member 21, LS, female)

The ability to work without the supervisor is perceived as the core requirement for independence. The reviewer argues that it is crucial to check if an applicant has developed a research topic independently from the supervisor’s. The interviewee argues that clear and hard criteria for assessing independence exist, although those are not explicitly specified by the ERC (except for “one publication without a supervisor”). It becomes evident that already one reviewer operates with a heterogeneous set of indicators to measure independence. This is even more the case when all reviewers in a panel bring in their individual understanding. Besides applying different indicators, indicators also are weighted differently. Research has shown that evaluators tend to overstate/understate the relevance of a criterion, depending on

whether it is attributed to a male or a female applicant (Uhlmann and Cohen 2005, Baltes and Parker 2000).

We found, that very different achievements are evaluated, because formalised indicators are not perceived as binding and indicators are applied individually. Therefore it remains unclear what is exactly measured by the independence criterion (intransitivity) and that it is applied unsystematically by panel members.

3 Independence: relevance and limitations

Next, we analyse how independence is constructed in panels, looking at the expectations and assumptions that are attributed to independence as a criterion for excellence. This relevance is discussed as well as potential limitations and concerns of this indicator.

Some reviewers assess independence as the core of excellence, as most important to produce new knowledge and advance science beyond its current frontiers.

But one has to bring it up all the time, this thing with independence. (...) I always bring it up, because I think we really need to encourage the younger researchers to go for independence and really distance themselves and go to a new level. (...) We need to get independence and new ideas and new ways in science. So I think this is something that one should always look at. (Panel member 12, LS, female)

Here, the construction of independence includes a social perspective. The reviewer argues that it is crucial for young researchers to develop their own research strategy in order to guarantee progress in science. Research funding might push young scientists in this direction.

Another reviewer points out that publishing with the supervisor indicates a lack of independence.

I always question independence if you still co-publish with your PhD or postdoc supervisor. (Panel member 19, LS, female)

A rather opposite position refers to limit the relevance of independence for various reasons. One reason relates to age and experience. Researchers at the beginning of their career, like Starting Grant applicants, are perceived as too young to prove independence. For them, independence is not seen as a given requirement, but as a future objective. This is in line with the formal ERC definition that grants should provide the time and resources to develop an independent research team or programme.

If it's a Starting Grant, a person cannot be expected to be independent on this point. You are funding him to become independent and strong (panel member 21, LS, female)

I think "independent" was also one criterion which I thought was odd because indeed these people come from their Postdoc, so one cannot yet evaluate them on their independence, in my opinion. One can say: "Ok, they have done great work as a Postdoc". And this would mean together with a supervisor. So it would be "participated in great work". (Panel member 6, LS, male)

As it is argued that young researchers are not yet able to demonstrate (topical) independence, it is suggested to evaluate not the demonstrated independence, but rather the extent to which they have participated in excellent research. But the challenge is to identify and evaluate the personal contribution of (young) researchers.

It is difficult, if they came from a big lab, you follow what they are doing there, so this is not creative thinking, maybe it was creative when you were there, but it is very difficult

to seek ... the influence of your mentor and what is your own. (Panel member 16, LS, female)

Another suggestion to make independence a more appropriate criterion for younger researchers' excellence is to evaluate the geographic independence rather than the topical or social dimension. Geographic mobility is perceived as sufficient prove for being on the way to independence. In reviewers' construction of independence hardly focuses on the research topic, but on the applicants' mobility between research institutions.

*I personally think it's not fair to do a Postdoc and be expected to start something totally new. I think if you're **in a new location** you'll **naturally diverge**. So I actually don't pay much attention to that – if you're in a new location. (Panel member 3, LS, female)*

And finally, a critical issue was mentioned that generally puts into question the need to prove (topical) independence: While nowadays research is complex and highly organised in teams, the need to prove independence might not reflect current research reality.

I don't think it's necessarily good globally, because the fact that they put a lot of emphasis on the contribution of one researcher ... Sometimes, there were discussions about "Is this really an individual idea, or the idea of a group?" and "What was his or her contribution to that?" and it was putting a lot of emphasis on individuals as researchers and not on individuals performing well or not with other researchers. (Panel member 24, LS, female)

Research argues that in average, women tend to emphasis the efforts and contributions of teams to their work more than men (Sarsons 2015). Women therefore might feel more reluctant to put themselves as single author on a publication and not contributing team members. In our online-survey, applicants have in fact claimed that the focus on proving independence gives a wrong impression of one's own merit.

*Having to sell this entire team project as if it was MY OWN, as if I was a "leader" when in fact I'm always **collaborating** and learning from everyone. I guess this is also a woman thing. It's really a perception. But an important one. I would have much rather said that x will do this, and y will do that, rather than "I will deliver x..., y..." (Applicant, female)*

In this applicant's understanding it would be fair to also name the contributions of her colleagues instead of presenting the projects as hers. Reviewers have also questioned the relevance of independence in contrast to collaboration as a further prerequisite for excellence:

I think independence sometimes is misunderstood; that independence creates the heart of all collaboration and interaction. I have been collaborating with my former boss all my life. We produced more than 100 papers together and now here he's part of my lab. So, I know what independence is and what collaboration is. Sometimes removing a collaboration effort with your former boss or seniors in your area is negative for the development of science and for your personal development. (Panel member 21, LS, female)

These quotes illustrate that proving one's independence by working without a supervisor may limit scientific improvements. It is argued that collaboration with the supervisor not only contributes to excellent scientific research, it also allows independence.

4 Gender practices of independence

In a next step, we take the applicants' sex into account and discuss if and in which way independence and appropriate indicators are deployed differently to female and male applicants. We found that topical/social independence as well as mobility (= geographic independence) are partly dropped for men. Interviewees reported that in some peer review panels, independence and mobility are checked more for female than for male applicants.

*Women don't move as soon and as long as men to another country to make part of their research there. At the same time I noticed that some **men have never moved out** of their university (...). They start their PhD at the same university. And they also become professors at the same university. And everybody finds that they have an excellent CV.*
(Panel member 32, SH, female)

This illustrates that mobility is well checked for female, but not necessarily for male applicants. The following quotes also show that indicators are applied differently to female and male applicants.

*"I think [panel members] bring up **independence** more as an issue with females and they **tend to oversee it for males**."* (Panel member 12, LS, female)

The one thing that I've experienced ... is that with this independence mark males can get away much easier than females. Females' independence is questioned more than males' is. (...) You're not less independent as a female, just because you still co-publish with either your PhD or Postdoc supervisor than if a male would do it. There is no difference. But that is clearly seen upon as differently. And the males come out much better than the females in that aspect. (Panel member 19, LS, female)

These panel members report that independence is questioned more for female applicants than for their male counterparts. Men who (still) co-publish with their supervisor are perceived as excellent, women who do the same are blamed for a lack of independence. Independence seems to be something that does not need to be questioned for men. The criterion is dropped, as men are perceived as "naturally" independent. This refers to gendered attribution and stereotypes and favours male applicants. It illustrates that when excellence is constructed, different standards to evaluate independence for female and male applicants are deployed in the assessment process. The same observation of shifting standards was also made by Ahlqvist et al. (2015).

5 Word use on independence

Can gendered practices also be observed in evaluation reports? Comparing word use for female versus male applicants, we found that independence is used as frequently for female applicants as for male applicants. Taking into account that women have more negation words in their reports, we would expect that words referring to independence are more often combined with negation words like no, non, hasn't, isn't ... in reviews of female applicants. Whether this gendered pattern can be observed will be investigated in a next step. (For details about the linguistic analysis of review reports see van den Besselaar, Stout, Gou in this proceedings volume.)

6 Gendered character of independence

Last, we analyse if independence as criterion for constructing excellence itself or the defined indicator (“to have at least one publication without the supervisor”) is gender biased. Gender bias is defined as prejudice for or against one gender relative to the other (Heilman 2012), and the term is commonly used as “bias against women”. Stereotypes are a structured set of beliefs about the attributes of a social group that are ascribed to members of that social group. Gender stereotypes refer to beliefs about what men and women *are* like (descriptive stereotypes) and *should be* like (prescriptive stereotypes). Often they are unconscious. So for instance men are stereotypically described as **independent** or rational and women as collaborative or understanding to women (Bakan 1966). Gender bias will occur when the attributes required for a particular position or role are more strongly associated with stereotypical attributes of men than of women (Heilman et al. 2015). As science is strongly associated with stereotypical attributes of men, women are not judged for what they are or have done, but for stereotypic conceptions of women in general and respective expectations. When fitting well to stereotypic feminine characteristics such as being collaborative or team-orientated (not self-promoting), women fit less to masculine characteristics like independence. Women are penalised for the attributes ascribed to them in general and at an individual level.

When studying how independence is constructed in ERC peer review panels, we found gender stereotypes that might explain gender practices described above. When talking about *independence from the PhD supervisor* female and male applicants are perceived clearly differently:

*In my experience, **women** are much more satisfied when they can **collaborate** with someone they know well, that gives safety. While **men have the ambition to kick off**, to start their own thing as early as possible.* (Panel member 31, SH, male)

This panel member portrays female applicants as appreciating to be dependent while male applicants are perceived as agentic and willing to become independent as soon as possible. It becomes almost evident that men’s independence is not questioned in the panel discussions. Rather it is assumed that all men are striving for independence. It seems that it is almost a personal game, as the next quote shows:

*The supervisor doesn’t want any competition. This competitiveness is still very dominant. And perhaps **women** think that they **don’t want to work against** their former boss. And **men think: “I will show him!”*** (Panel member 1, LS, female)

In this context, the postdoc researcher is supposed to be brave to challenge the supervisor, which in the panel member’s point of view is less common for women. Women are portrayed as more modest and less willing to be in direct competition, whereas men are. This is closely linked to independence as a criterion for excellence as this is more naturally attributed to male researchers which disadvantages women.

They were women and maybe they were less good at arguing with their supervisor and saying “I need to look more independent. I need my name alone on my paper. Nobody’s going to believe me if my name isn’t alone. (Panel member 3, LS, female)

Here, the applicant’s position in the research group and the personal relation to the supervisor are linked to the scientific outcome. This is related to feminine attributions like the lack of being assertive and insisting on being the lone author. These personal characteristics seem decisive when excellence is assessed, as another panel member lines out:

Women are perhaps less brutal negotiating [for independence]. (Panel member 1, LS, female)

It needs to be pointed out that gendered assumptions come from panel members of both sexes:

Several of the women had stayed with their advisors. And their advisors were saying: "They are going to be totally independent" and I didn't believe it. (Panel member 7, LS, female)

By the list of quotes we intend to demonstrate that gender stereotypes exist and that they do influence the assessment of independence in peer review panels. As independence and mobility are more attributed to male applicants and their actual independence is less scrutinised by panel members. So different standards are deployed, based on existing (unconscious) gender stereotypes.

7 Conclusions

In this paper, we have discussed independence as an indicator for measuring scientific excellence in the ERC peer review process and analysed how independence is formalised at ERC level. We found that there is only one indicator ("to have one publication without the supervisor") formally defined. Another part of the formalisation ("to develop an independent research team") is ignored. Both, criterion and indicator, might be gendered as the assessment is based on gender stereotypes. On the level of review panels, we have demonstrated that the indicator "independence" is deployed differently to female and male applicants: Panel members assume that male applicants are independent and mobile (geographically independent) whereas women in general terms are not. Therefore, in some panels, independence is not questioned for male applicants (it is taken for granted). Male applicants who are in fact not independent or not mobile benefit from this assumption as the criterion is dropped. Contrary, women as a group are - based on stereotypical assumptions - expected to be less independent. So female applicants are checked more thoroughly for their independence and it is therefore detected more frequently as a shortcoming in terms of excellence. This stereotypical assumption about the social group of women is disadvantaging individual female applicants.

Furthermore, we found that – due to a lack of clear definition of the criterion and appropriate indicators – independence is constructed and applied unsystematically in peer review panels. This demonstrates that all prerequisites for more transparency and quality control of "classical" excellence indicators also need to be applied for them.

In an evaluation process, the discrepancy between (male) attributes perceived as important for success and female stereotypes produce negative performance expectations that lead to gender biased judgements – disfavours women. That can best be avoided by clearly defined evaluation criteria: *"The more vague and poorly defined the judgement criteria, the more easily information can be distorted to fit expectations."* (Heilman 2012, 119). Therefore, criteria should be specified more in detail, processes should be stronger standardized, and panel members should be made aware of how gender bias emerges in assessment practices.

At the same time, these heterogeneous constructions illustrate the broad potential this indicator covers: It touches the question of originality and collaboration, of geographic mobility as well as of power structures within research teams.

Further research is needed to get better insights in the relevance of this indicator. To learn about topical independence or cognitive distance will also be a rewarding step to better understand how ground-breaking research is assessed.

REFERENCES

- Ahlqvist, V., Andersson, J., Söderqvist, L., Tumpane J. (2015). A gender neutral process? – A qualitative study of the evaluation of research grant applications 2014, Swedish Research Council, Stockholm.
- Aksnes, D. W., Rorstad, K., Piro, F., and Sivertsen, G. (2011). Are female researchers less cited? A large-scale study of Norwegian scientists. *Journal of the American Society for Information Science and Technology*, 62(4), 628–636.
- Bakan, D. (1966) *The duality of human existence*, Chicago.
- Baltes, B. B. and Parker, C. P. (2000) Reducing the effects of performance expectations on behavioral ratings, in: *organisational Behavior and Human Decision Processes* 82(2) 237–267.
- Bleijenbergh, I.L., van Engen, M.L. u. Vinkenburg, C.J. (2012). “Othering women. Fluid images of the ideal academic”, *Equality, Diversity and Inclusion: An International Journal*, 32(1), 22–35.
- Brouns, M. and Addis, E. (2004). Synthesis report on the workshop. European Commission (2004): *Gender and Excellence in the Making*.
- EC European Commission (2013). *ERC Work Programme 2014*.
- EC European Commission (2014). *Self-Evaluation form. ERC Starting Grant and Consolidator Grant 2014*
- Heilman, M. E., Manzi, F., Braun, S. (2015). Presumed incompetent: perceived lack of fit and gender bias in recruitment and selection. In: A. M. Broadbridge and S. L. Fielden (Eds.) *Handbook of gendered careers in management. Getting in, getting on, getting out*, 90–104, Cheltenham, Edward Elgar.
- Heilman, M.E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior*, 32, 113–135.
- Lamont, M. (2009). *How professors think: inside the curious world of academic judgment*, Harvard College.
- O’Connor, P. and O’ Hagan, C. (2015). Excellence in university academic staff evaluation: a problematic reality? *Studies in Higher Education*, 1–15.
- Rees, T. (2011). The Gendered Construction of Scientific Excellence. *Interdisciplinary Science Reviews*, 36(2), 133–45.
- Sarsons, H. (2015). *Gender Differences in Recognition for Group Work*, Working Paper.
- Uhlmann, E. L. and Cohen, G. L. (2005) Constructed criteria: Redefining merit to justify discrimination. *Psychological Science*, (16), 474–480.
- Van den Besselaar, P., Sandström, U. van der Weijden, I. (2012). *The Independence Indicator: Towards Bibliometric Quality Indicators at the Individual Level*
- Van den Brink, M. and Benschop, Y. (2012). Gender practices in the construction of academic excellence: Sheep with five legs. *Organization*, 19(4), 507–524.
-

CHAPTER 11

Social Sciences and Humanities



Indicators for Research Performance in the Humanities? The Scholars' View on Research Quality and Indicators¹

Michael Ochsner^{*}, Sven E. Hug^{**}

*ochsner@gess.ethz.ch

Department of Humanities, Social and Political Sciences, ETH Zurich, Muhlegasse 21, 8001 Zurich (Switzerland) and FORS, University of Lausanne, Géopolis, 1015 Lausanne (Switzerland)

**sven.hug@gess.ethz.ch

Department of Humanities, Social and Political Sciences, ETH Zurich, Muhlegasse 21, 8001 Zurich (Switzerland) and Evaluation Office, University of Zurich Muhlegasse 21, 8001 Zurich (Switzerland)

INTRODUCTION

In this paper, we present indicators for research quality in the humanities collected in our previous work (Ochsner, Hug, & Daniel, 2012). We focus on how these indicators are accepted by humanities scholars. We also investigate differences between different subgroups of the humanities scholars we surveyed with regard to their preferences for such indicators.

We address the theme of the conference ('peripheries frontiers and beyond') regarding two notions of (scientometric) periphery: First, we investigate indicators for research quality in the humanities, a field where there is a lack on knowledge on how to assess or even measure research quality, in fact, there is a quite broad consensus that (evaluative) bibliometrics cannot be readily applied in the humanities (Hicks, 2004; Nederhof, 2006). Second, we fully cover three humanities disciplines at Swiss universities and member universities of the League of European Research Universities (LERU). Scholars are a neglected stakeholder when it comes to the design of research assessment procedures or the selection of research indicators. However, they are directly affected, they know best what research quality in their field is and what impact the use of certain indicators could have on their research practices.

The paper is structured as follows: first, we present the background for selecting indicators for research quality. This is followed by a description of our analysis methods and the presentation of the results. We finally discuss the results with regard to their use in research evaluation.

INDICATORS FOR RESEARCH PERFORMANCE LINKED TO QUALITY CRITERIA

Humanities scholars have many objections against research evaluation, especially against quantification of research performance. This is at least partly due to the fact that there is a missing link between indicators for research performance and research quality (Ochsner et al., 2012). Scientometricians also note that research indicators are only loosely tied to quality definitions (Brooks, 2005; Donovan, 2008). Such weak or missing links between indicators

¹ This work was supported by the Rectors' Conference of the Swiss Universities (CRUS) within the framework of the SUK B-05 Innovation and Cooperation Project 'Mesurer les performances de la recherche' (Measuring Research Performance) as part of the cooperative initiative of the Universities of Zurich and Basel entitled 'Developing and Testing Research Quality Criteria in the Humanities, with an Emphasis on Literature Studies and Art History'. Matching funds for the initiative were provided by the University of Zurich.

and quality make it difficult for the assessed scholars to understand what is being measured. Therefore, the reluctance of humanities scholars to accept a quantitative representation of research quality is not surprising. At the same time, if the measurement is not or only loosely tied to the object that is to be measured (i.e. research quality), unintended effects become more likely. A sound measurement approach can replace the missing links between indicators and the concept(s). This means that before one can measure a concept with indicators, the concept needs to be clearly defined (Lazarsfeld & Barton, 1951, p. 155). Borsboom, Mellenberg, and van Heerden (2004, p. 1067) formulate the need of the definition of the concept in the following way: '[The issue is not] first to measure and then to find out what it is that is being measured but rather that the process must run the other way'. In a project on research quality and assessment in the humanities, we applied this approach by defining our concept ('research quality') by explicating quality criteria (Hug, Ochsner, & Daniel, 2014). In a next step, every quality criterion is specified and defined explicitly by one or more aspects (i.e. the analytical definition). Then, each aspect is operationalized by one or more indicators that specify how the aspect can be observed, quantified or measured (i.e. operational definition). Of course it is possible that for some aspects no indicators can be found, thus such an aspect cannot be measured by indicators.

Using Repertory Grid interviews (Ochsner, Hug, & Daniel, 2013) and a Delphi survey (Hug, Ochsner, & Daniel, 2013), we found 19 criteria for research quality in the humanities, specified by 70 aspects. We then identified aspects that reach a consensus among the humanities scholars. These aspects can be used to assess research quality in the three disciplines we studied (German and English literature studies and art history). In a next step, we collected indicators for research quality from the literature and directly from humanities scholars during the Repertory Grid interviews and the Delphi survey. This resulted in a long list of indicators, some very specific, some very vague. We grouped them into 62 indicator groups and linked them to the quality aspects they can potentially measure (for a complete list, see Ochsner, et al., 2012). Humanities scholars then rated the indicators according to their utility in measuring the corresponding quality aspects. In this paper, we will investigate differences in preferences for indicators between subgroups of our population.

METHOD

We designed a questionnaire to rate the indicator groups linked to quality aspects.² The scholars had to rate the indicator groups according to a statement on a 6-point scale (1 'strongly disagree' to 6 'strongly agree'). The statement consisted of two parts: A generic part ('The following quantitative statements provide peers with good indications of whether I...') followed by an aspect (e.g. 'I participate in a scholarly discourse regarding my field') of a criterion (e.g., 'scholarly exchange'). The scholars were presented the indicator groups that can potentially measure the given aspect and they had to rate each indicator group assigned to this aspect according to the statement. Because there were some discipline-specific aspects (i.e., aspects that reached consensus only in one or two disciplines), the questionnaires differed between the three disciplines. In German literature studies (GLS), the scholars had to rate 86 items consisting of 59 unique indicator groups assigned to 19 aspects (some indicator groups can be assigned to more than one aspect). In English literature studies (ELS), the respondents had to rate 85 items consisting of 45 unique indicator groups, and in art history

² For example, scholars were asked to indicate their (dis-)agreement with indicators as follows: The following quantitative statements provide peers with good indications of whether I participate in a scholarly discourse regarding my field: (a) number and weighting of publications for a disciplinary audience, (b) number of sources from my discipline I quote in my publications, (c) number, weightings and durations of editorships in my discipline etc.

(AH), the scholars had to rate 74 items consisting of 44 unique indicator groups assigned to 15 aspects. The questionnaire was administered in German and in English. Invitations were sent to all scholars holding at least a PhD working in one of the three disciplines at a Swiss university or at a member university of the League of European Research University (LERU). All in all 664 invitations were sent out. The field period lasted from October 2011 to January 2012.

We analyse the data using descriptive statistics such as means and medians to describe the acceptance of the indicators to measure the relevant quality aspects in the three disciplines. In this paper, we focus only on those indicators that have been rated in all disciplines (i.e. indicators measuring quality aspects reaching consensus in all disciplines) because the goal is to compare between different subgroups of the sample, including discipline. We also identify indicator groups that reach consensus. We define consensus the same way as we defined it concerning quality aspects, i.e. the median is above 4 (50% of the scholars rated the indicator group with a 5 at least) and the 10th percentile is above 3 (not more than 10% of the scholars reject the indicator group). Because we did not use a random sample but a population survey, we cannot use inferential statistics. Therefore, we use bootstrap resampling (with 1000 replications) to estimate the stability of the results (95%-'stability intervals', see, e.g., Schneider & van Leeuwen, 2014) and standardized effect sizes to analyse differences in means across subgroups.

RESULTS

In total, 133 out of 664 questionnaires have been returned which corresponds to an overall response rate of twenty per cent. Among the respondents were 48 scholars of GLS, 43 scholars of ELS, and 42 scholars of AH (corresponding to response rates of 23%, 22%, and 17% respectively). Fifty-two respondents were members of Swiss universities and 81 respondents were members of LERU universities (corresponding to response rates of 33% and 16% respectively). Fifty-six women and 77 men (corresponding to response rates of 21% and 19% respectively) participated in the survey. Because the questionnaires differed between disciplines, an analysis of all indicator groups can only be carried out by discipline (which does make sense as we are looking for indicators that adequately inform on quality criteria in a discipline). Most indicator groups were accepted by a majority of our respondents if analysed per discipline (acceptance being defined as a median higher than '4'). In GLS, 93% of the items reached that threshold, in ELS, 91% and in AH 97% respectively. However, also a minority that is not to be neglected clearly disagreed with many indicators: only 10 indicator groups (12%) reached consensus in GLS, one indicator group (1%) in ELS, and 16 indicator groups (22%) in AH. For a more information on the results of the whole questionnaire, see Ochsner, Hug, & Daniel, 2014).

In this paper, we focus on those 39 items that have been part of all three questionnaires in order to investigate whether there are differences between different subgroups of our sample. The 39 items consist of 34 unique indicator groups assigned to 8 quality aspects specifying 7 quality criteria. Of the 39 items, three indicator groups (8%) were rejected by a majority. However, only two indicator groups (5%) reached consensus over all respondents (see table 1). If we look at mean differences between disciplines, we see that most differences are small to moderate (*Cohen's d* < 0.8). However, we find also that ELS scholars rated the indicators quite lower than GLS and AH scholars (8 items with a *Cohen's d* > 0.5 in ELS vs. GLS, 18 in ELS vs. AH) and AH scholars rated some items higher than GLS scholars (4 items with a *Cohen's d* > 0.5). Regarding gender, there are no big differences in means, only 8 items exhibit

a *Cohen's d* between 0.2 and 0.3, which can be considered small. We find some differences, however, between tenured and non-tenured scholars: tenured are more in favour of the indicator group 'initiation/foundation' (number and weighting of what the person has initiated or founded, e.g. book series, institutions, journals etc.), no matter whether

Table 1: Overall Mean, percentage of negative ratings (bootstrapped 95% stability intervals in parentheses), and Cohen's *d* of subgroups for indicator groups.

Variable	Mean	% of negative ratings	Cohen's <i>d</i> GLS vs ELS	Cohen's <i>d</i> AH vs GLS	Cohen's <i>d</i> ELS vs AH	Cohen's <i>d</i> Gender	Cohen's <i>d</i> Tenure
Publications: disciplinary exchange	4.95 (4.77-5.14)	0.06 (0.02-0.10)	0.22	0.27	0.27	-0.04	0.02
References: disciplinary exchange	3.86 (3.62-4.10)	0.30 (0.22-0.38)	-0.15	0.65	0.65	0.02	0.09
Presentations: disciplinary exchange	4.55 (4.35-4.75)	0.14 (0.08-0.19)	0.24	0.00	0.00	0.11	0.04
Editorship: disciplinary exchange	4.35 (4.13-4.56)	0.21 (0.14-0.28)	0.26	-0.33	-0.33	-0.10	0.16
Organized events disciplinary exchange	4.44 (4.24-4.65)	0.17 (0.10-0.23)	0.51	0.02	0.02	0.03	-0.16
Collaborations: disciplinary exchange	4.35 (4.13-4.58)	0.19 (0.12-0.25)	0.83	-0.06	-0.06	-0.04	-0.12
Personal contacts: disciplinary exchange	3.92 (3.68-4.17)	0.30 (0.22-0.38)	0.25	0.05	0.05	0.29	0.15
Review Activities: disciplinary exchange	4.20 (3.99-4.41)	0.20 (0.13-0.27)	0.20	0.12	0.12	-0.06	0.37
Academic associations: disciplinary exchange	3.98 (3.78-4.18)	0.28 (0.20-0.36)	0.09	0.04	0.04	0.27	0.14
Panels: disciplinary exchange	4.14 (3.94-4.35)	0.23 (0.15-0.30)	0.22	0.07	0.07	0.10	0.22
Survey: renewal of interpretations of the past	3.81 (3.59-4.04)	0.35 (0.27-0.43)	0.19	0.85	0.85	-0.12	-0.05
Citations: impact on research community	3.74 (3.49-3.98)	0.35 (0.27-0.43)	-0.05	0.05	0.05	0.08	-0.01
Acknowledgements: impact on research community	3.20 (2.94-3.45)	0.54 (0.46-0.62)	-0.42	0.19	0.19	0.15	0.24
Success of junior researchers: impact on research community	4.27 (4.07-4.48)	0.21 (0.14-0.28)	0.15	0.21	0.21	0.07	0.27
Started initiatives: impact on research community	4.02 (3.78-4.25)	0.30 (0.22-0.38)	-0.09	0.07	0.07	0.09	0.54
Editorship: impact on research community	4.04 (3.81-4.26)	0.25 (0.17-0.32)	0.46	-0.24	-0.24	0.00	0.25
Opportunities for junior researchers: openness to other persons	4.47 (4.25-4.69)	0.18 (0.12-0.25)	0.16	0.40	0.40	-0.20	-0.02
Assessed openness: openness to other persons	4.81 (4.62-5.00)	0.08 (0.04-0.13)	0.27	0.21	0.21	-0.15	-0.03
Heterogeneity of junior researchers: openness to other persons	4.20 (3.97-4.42)	0.26 (0.18-0.33)	0.20	0.41	0.41	-0.07	-0.33

Table 1: continued

Variable	Mean	% of negative ratings	Cohen's d GLS vs ELS	Cohen's d AH vs GLS	Cohen's d ELS vs AH	Cohen's d Gender	Cohen's d Tenure
Assistance: openness to other persons	3.62 (3.39-3.85)	0.41 (0.32-0.49)	0.19	0.23	0.23	0.24	-0.15
Course accessibility: openness to other persons	4.23 (4.00-4.47)	0.27 (0.19-0.35)	0.52	0.14	0.14	0.11	-0.34
Availability of publications: openness to other persons	3.95 (3.71-4.20)	0.35 (0.27-0.42)	0.06	0.39	0.39	0.11	-0.11
Sources: rich experience with sources	4.67 (4.43-4.90)	0.19 (0.12-0.25)	0.11	0.23	0.23	0.33	0.18
Documentation activities: rich experience with sources	3.98 (3.74-4.23)	0.33 (0.25-0.41)	0.54	0.24	0.24	-0.05	-0.08
Output of documentation activities: rich experience with sources	4.14 (3.91-4.36)	0.27 (0.20-0.34)	0.54	0.21	0.21	-0.06	-0.09
Research time: knowledge based on own research	4.32 (4.04-4.59)	0.25 (0.18-0.32)	0.40	0.02	0.02	-0.09	-0.14
Personal library: knowledge based on own research	3.14 (2.85-3.43)	0.57 (0.48-0.66)	-0.19	0.52	0.52	0.02	0.25
References: knowledge based on own research	3.81 (3.56-4.07)	0.38 (0.30-0.46)	0.22	0.29	0.29	0.22	0.03
Monographs: knowledge based on own research	4.26 (4.04-4.49)	0.22 (0.15-0.29)	0.13	0.25	0.25	-0.02	0.11
Attractivity to junior researchers: arouse passion for research	4.59 (4.41-4.78)	0.13 (0.07-0.18)	0.04	0.21	0.21	-0.06	0.00
Qualification of junior researchers: arouse passion for research	4.53 (4.33-4.72)	0.17 (0.10-0.23)	0.21	0.33	0.33	0.04	0.06
Success of junior researchers: arouse passion for research	4.35 (4.13-4.58)	0.23 (0.16-0.31)	0.20	0.24	0.24	-0.04	0.07
Teaching awards: arouse passion for research	3.66 (3.44-3.88)	0.40 (0.31-0.48)	0.20	-0.23	-0.23	0.14	0.12
Acknowledgments: arouse passion for research	3.19 (2.96-3.42)	0.56 (0.47-0.64)	-0.32	0.18	0.18	0.27	0.25
Survey enthusiasm teaching: arouse passion for research	4.15 (3.94-4.36)	0.26 (0.18-0.33)	0.30	0.01	0.01	0.21	-0.27
Survey enthusiasm public: arouse passion for research	3.38 (3.15-3.62)	0.50 (0.41-0.58)	0.40	0.27	0.27	-0.16	-0.18
Started initiatives: pointing out important research for the future	3.84 (3.61-4.08)	0.32 (0.24-0.40)	-0.31	0.07	0.07	0.17	0.45
Strategies: pointing out important research for the future	4.00 (3.77-4.23)	0.34 (0.26-0.42)	0.28	0.12	0.12	-0.04	0.13
Utilizing sources: pointing out important research for the future	4.03 (3.81-4.25)	0.29 (0.21-0.36)	0.59	0.18	0.18	-0.17	-0.19

it measures the quality aspect 'impact on research community' or 'vision of the future'. It is striking that among the items with a *Cohen's d* above 0.2, those which presuppose a strong network (reviews, board memberships, acknowledgements, success of young scholars,

initiations/foundations, editorships, personal library) are rated more favourably by tenured scholars while those which refer to teaching, collaboration and social competence (heterogeneity of students and staff, openness and accessibility to courses, survey of students and junior researchers on arousing passion for the subject) are rated more favourably by non-tenured scholars.

To be effective, indicators applied in research evaluation need to be accepted by those who are directly affected by them (i.e. the scholars). Therefore, indicators with a high degree of acceptance, i.e. a consensus, need to be identified. In our definition, an indicator reaches consensus if less than 10% of the scholars rate it negatively (i.e., with 1, 2 or 3) and if, at the same time, 50% or more rate it with 5 or 6. Only two out of the 39 items reach consensus over all respondents (*publications* measuring ‘disciplinary exchange’ and *assessed openness* measuring ‘openness to other persons’). With regard to consensus, we find some differences between disciplines. While in AH 8 items and in GLS 4 items reach consensus, none of the items reaches consensus in ELS. Only two items reach consensus in two disciplines: not surprisingly the ones that reach consensus over all respondents. These are also the two items attracting the smallest proportion of ELS scholars rating them negatively (16% for both items). If we look at the stability intervals derived from bootstrap resampling, however, no item has a stability interval that does not include values above 10% (of persons choosing a negative rating), one item just scratching the threshold with the upper level of the stability interval being only slightly higher than 10% (*publications* measuring ‘disciplinary exchange’). With regard to disciplines, two items have stability intervals that do not exceed the 10% threshold in AH (in fact, for these two items, not a single respondent in AH chose a negative value: *publications* measuring ‘disciplinary exchange’ and *assessed openness* measuring ‘openness to other persons’), in GLS one item shows a stable consensus (*publications* measuring ‘disciplinary exchange’) and in ELS none.

If we compare the ratings by gender, we find only one item that reaches consensus among men and women (*publications* measuring ‘disciplinary exchange’) and two more reaching consensus among men (*assessed openness* measuring ‘openness to other persons’ and *attractivity to young researchers* measuring ‘arouse passion for research’). However, no item reaches a stable consensus, yet two items that only reach consensus among men miss the 10% threshold only slightly, thus could be considered as stable (10.05% and 10.3% resp., again the *publications* and *assessed openness* we know from the other comparisons). Similarly, there are not many differences between tenured and non-tenured scholars: the *publications* measuring ‘disciplinary exchange’ reach consensus among both tenured and non-tenured scholars (the upper limits of the stability intervals being 13% for tenured and 10.3% for non-tenured, i.e. quite stable consensus). All other items do not reach consensus except *assessed openness* measuring ‘openness to other person’ that reaches consensus only among tenured scholars (the upper level of the stability interval being 11%).

CONCLUSION

This paper examines the scholars’ acceptance of indicators to measure research quality in the humanities. In our previous research, we already found that humanities scholars are open regarding research evaluation using quality criteria that relate to their own notions of quality but are reluctant to accept indicators or a quantitative approach towards research evaluation (Ochsner et al., 2014). We also found that there is a mismatch of quality criteria between research evaluators and humanities scholars (Hug et al., 2013). In this paper, we investigated whether there are differences in preferences for research indicators among subgroups of

humanities scholars. While we found small differences between the three disciplines we studied, we did find only few differences between other subgroups (gender and tenure). Our analysis shows that while most indicators would be accepted by at least 50% of our respondents³ in an informed peer review process, almost all indicators face rejection from a large minority (regarding the indicators in this study on average 28%). The only indicators that reached consensus in all three disciplines were *publications* measuring ‘disciplinary exchange’ and *assessed openness* measuring ‘openness to other persons’. From this, it follows that a purely quantitative approach to research evaluation is rejected by a vast majority of our respondents.

We can conclude that the use of quantitative information in the evaluation of humanities research is possible if some restrictions are considered. The indicators must be linked to the humanities scholars’ quality notions. They also have to be accepted by the scholars in order not to interfere in a destructive way with their research practices. While many scholars agree to some indicators measuring certain quality criteria during an informed peer review process, our results suggest that the use of the indicators should be agreed upon with the scholars. We found that irrespective of gender, tenure, and discipline, a fairly large part of the scholars oppose most indicators. However, we have chosen three ‘aesthetic’ disciplines considered especially difficult to evaluate in a quantitative way. It is likely that in more ‘empirical’ disciplines, the use of a broader set of indicators will be accepted. Nevertheless, research evaluation in the humanities, especially quantitative measurements, should be discursive as well as participatory and should focus on research quality or at least include it to an important degree.

REFERENCES

- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, *111*(4), 1061–1071. doi:10.1037/0033-295X.111.4.1061.
- Brooks, R. L. (2005). Measuring university quality. *The Review of Higher Education*, *29*(1), 1–21. doi:10.1353/rhe.2005.0061.
- Donovan, C. (2007). The qualitative future of research evaluation. *Science and Public Policy*, *34*(8), 585–597. doi:10.3152/030234207X256538.
- Hicks, D. (2004). The four literatures of social science. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of quantitative science and technology research: The use of publication and patent statistics in studies of S&T systems* (pp. 476–496). Dordrecht: Kluwer Academic Publishers.
- Hug, S. E., Ochsner, M., & Daniel, H.-D. (2013). Criteria for assessing research quality in the humanities: A Delphi study among scholars of English literature, German literature and art history. *Research Evaluation*, *22*(5), 369–383. doi:10.1093/reseval/rvt008.
- Hug, S. E., Ochsner, M., & Daniel, H.-D. (2014). A framework to explore and develop criteria for assessing research quality in the humanities. *International Journal for Education Law and Policy*, *10*(1), 55–64.
- Lazarsfeld, P. F., & Barton, A. H. (1951). Qualitative Measurement in the Social Sciences. Classification, Typologies, and Indices, in D. Lerner & H. D. Lasswell (Eds.), *The Policy Sciences* (pp. 155–192). Stanford, CA: Stanford University Press.

³ From our previous studies with the same panel, we assume that there might be a selection bias regarding respondents being more open to indicators than non-respondents.

- Nederhof, A. J. (2006). Bibliometric monitoring of research performance in the social sciences and the humanities: a review. *Scientometrics*, 66(1), 81–100. doi:10.1007/s11192-006-0007-2.
- Ochsner, M., Hug, S. E., & Daniel, H.-D. (2012). Indicators for research quality in the humanities: opportunities and limitations. *Bibliometrie – Praxis und Forschung*, 1(4). Retrieved from <http://www.bibliometrie-pf.de/article/view/157/192>.
- Ochsner, M., Hug, S. E., & Daniel, H.-D. (2013). Four types of research in the humanities: Setting the stage for research quality criteria in the humanities. *Research Evaluation*, 22(4), 79–92. doi:10.1093/reseval/rvs039.
- Ochsner, M., Hug, S. E., & Daniel, H.-D. (2014). Setting the stage for the assessment of research quality in the humanities: Consolidating the results of four empirical studies. *Zeitschrift für Erziehungswissenschaft*, 17(6 Supplement), 111–132. doi:10.1007/s11618-014-0576-4.
- Schneider, J. W., & van Leeuwen, T. N. (2014). Analysing robustness and uncertainty levels of bibliometric performance statistics supporting science policy. A case study evaluating Danish postdoctoral funding. *Research Evaluation*, 23(4), 285–297. doi:10.1093/reseval/rvu016.



Quality criteria and indicators for research in Theology –

What to do with quantitative measures?

Author 1:

Dr. Silvia Martens

Academic staff member Academic Services

T +41 41 229 51 25

silvia.martens@unilu.ch

University of Lucerne, Frohburgstrasse 3, P.O. Box 4466, 6002 Lucerne, Switzerland

Author 2:

Dr. Wolfgang Schatz

Academic Director

T +41 41 229 50 07

wolfgang.schatz@unilu.ch

University of Lucerne, Frohburgstrasse 3, P.O. Box 4466, 6002 Lucerne, Switzerland

ABSTRACT

In the work in progress report we would like to present preliminary results of our research project “Resource-based instrument for describing and evaluating research in the humanities and the social sciences as exemplified by theology”. Using a bottom-up approach that strongly involves the researchers we have worked to define criteria that are fit to adequately describe research (mapping of research activities, description of characteristics of the discipline) and evaluate research quality in theology.

We present the criteria set that we have thus developed and discuss the most likely areas as well as the challenges of its future application. One challenge in the application of the criteria set is certainly the diversity within the field of theology. Because theologians try to answer different research questions, use different research methodologies, have different audiences and intended research outcomes, researchers from the different theological sub-disciplines (exegetical theology, historical theology, systematic theology, practical theology) judge the importance of a part of the criteria differently. We also find differences between researchers from catholic and protestant faculties. Furthermore, while the majority of the – rather abstract – criteria were found suitable by the researchers, there is a reluctance to further specify them (in the form of indicators) for concrete scenarios of application and a major aversion to quantitative measures per se.

We present preliminary results of our research project “Resource-based instrument for describing and evaluating research in the humanities and the social sciences as exemplified by theology”. The project is one of several initiatives of the Swiss University Conference programme 2013–2016 P–3 « Performances de la recherche en sciences humaines et sociales » (Research output in the humanities and social sciences). The programme was launched to explore new ways to adequately compile information and assess the quality of research in the humanities and social sciences (HSS) and to enhance the visibility of research in HSS-disciplines.¹

Our project on theology aims to describe and visualise research in theology and identify quality criteria for this discipline.

Theology is a very interesting discipline for such a study because it has a multidisciplinary character with many overlaps to other disciplines like law, philosophy, religious studies historical studies, classical and ancient studies or social sciences and thus findings could probably be adapted to these disciplines. Otherwise we also consider the influence of non-academic actors like the churches on conceptions of quality in theological research.

To make sure that the quality criteria and the derived indicators are accepted in the discipline, the project follows a bottom-up approach that strongly involves the researchers. We conducted expert-interviews and focus group discussions as well as a 3-part online survey; we regularly consult with experts from theology engaging in an exchange with deans and faculty members. Additionally we studied the literature about research evaluation and observed current evaluation practices in the field (e.g. assessments of research proposals, criteria in nomination committees).

We have worked to define criteria that are fit to adequately describe research (mapping of research activities, description of characteristics of the discipline) and evaluate research quality in theology. These criteria may be applied to various dimensions of research that are relevant for different scenarios of analysis (i.e., evaluation of grant proposals, appointment procedures, evaluation and comparison of research institutions). Our criteria set includes core quality criteria on the following dimensions:

- the conception of a research project/research design (criteria applicable to research proposals but also retrospective evaluation),
- orientation of research (characteristics to describe research profiles rather than assess their quality),
- reception of research (as indirect or non-causal quality criterion),
- research performance (previous research output),
- competencies of the researcher (no criteria in the strict sense but may be consulted as indication of research qualification),

- research environment (criteria to assess research conditions of research institutions).

We propose core quality criteria for each of the above dimensions. Each of the core quality criteria in turn comprises several sub-criteria. Furthermore, in a second phase we have tried to identify indicators that can be applied to the rather abstract criteria when implemented in a specific evaluation process. On the basis of those criteria and indicators that have been accepted by the research community we intend to 1. formulate recommendations for evaluation of theological research and 2. create and provide an open-source software application to visualise research in theology. Both the recommendations and the software application may be further developed and adapted to other HSS disciplines in Switzerland.

In the work in progress report at the STI Conference we would like to present our criteria set and discuss the most likely areas as well as the challenges of its future application. One challenge in the application of the criteria set is certainly the diversity within the field of theology. Because theologians try to answer different research questions, use different research methodologies, have different audiences and intended research outcomes, researchers from the different theological sub-disciplines (exegetical theology, historical theology, systematic theology, practical theology) judge the importance of a part of the criteria differently. We also find differences between researchers from catholic and protestant faculties. Furthermore, while the majority of the – rather abstract – criteria were found suitable by the researchers, there is a reluctance to further specify them (in the form of indicators) for concrete scenarios of application and a major aversion to quantitative measures per se.

¹ For more information on the programme visit the website: <http://www.performances-recherche.ch> .



Social Impact Open Repository (SIOR). Transforming the peripheral space of social impact of research¹

Mar Joanpere* Elvira Samano** and Aitor Gómez

*mar.joanpere@ub.edu

CREA, University of Barcelona, Department of Social Theory, Av. Diagonal 690, 08034 Barcelona (Spain)

**elvira.samano.mx@gmail.com

University of Barcelona (Spain)

INTRODUCTION

The EC FP7 project “IMPACT-EV, Evaluating the impact and outcomes of EU SSH research” (2014-2017) aims at developing a permanent system of selection, monitoring and evaluation of the various impacts of Social Sciences and the Humanities research, with a very special attention to the social impact of research. The Work Package entitled “Identifying social impact of SSH research projects” has the main aim of analysing the social impact of SSH research and the factors that have contributed to obtain or not this impact, in order to create indicators to identify and evaluate the social impact of the SSH research ex-ante and ex-post.

This responds to a powerful trend that claims for the accountability of research in terms of social impact, as the use and the returns for the society of the investment that they have done to the researchers. This is clearly shown in the Reference Excellence Framework developed in UK, which establishes that a substantial part of the funding for the universities depend on the social impact demonstrated (in REF2014 a weighting of 20 per cent, and it is foreseen to be a 25 per cent of the funding). In many other countries and in the international arena, policy makers and the society are concerned with this social impact.

However, the measurement and indicators of these other impacts are less developed. In fact, as a more novel issue at the international arena, different models of assessing social impact with different roles of metrics are in discussion.

Importantly, Open Access has also a very big potential to change the relationship between science and society. As we now, the culture of Open Access is revolutionising the practices of research dissemination and discussion. Not only OA journals but also initiatives such as CORDIS or CRIS allow the scientific community, funding agencies and the whole society to find the records and main academic achievements of the researchers.

In this context, this paper explains the work in progress under the IMPACT-EV project to define indicators of Social Impact and a tool for displaying the social outcomes of research to the whole society.

¹ The research leading to these results has received funding from the European Union’s Seventh Framework Programme (FP7/2007- 2013) under Grant Agreement no 613202.

METHODS

The IMPACT-EV Work Package 6 is entirely dedicated to identifying methods and tools to conduct social impact assessments. This WP will develop case study research of success stories (projects which have achieved social improvements due to the implementation of policies or actions that are grounded in evidence from them). The selection of the success stories is based on the Ex-post evaluation of FP7 funded projects in the field of SSH, already developed in IMPACT-EV. The communicative evaluation of the social impact (CESI) of the success stories poses special relevance to the dialogue between researchers with end-users and stakeholders in the assessment of the actual social improvements related to the research outcomes. For the study of these cases, the consortium is conducting documental analysis and qualitative fieldwork (interviews and communicative daily life stories to researchers, end users and stakeholders). This WP will also identify how SSH research outputs are disseminated through the Web and on ways that increase social improvements. Finally, different working groups will be launched to study different challenges posed for the systematization of the information regarding social impact of SSH in order to contribute to the lac of indicators of social impact of research.

In addition, transversal activities are aimed to contribute to the definition of a sustainable and feasible system for SSH research evaluation at European and at national level. Literature review, comparative reports, integrative panels of experts and a Policy and Social dialogue group, and piloting evaluation criteria with national agencies are among these activities.

EXPECTED OUTCOMES

The research team has developed the Social Impact Open Repository (SIOR) which consists in an unprecedented data source at international level in which researchers display, cite and store the social impact of their research results (Flecha, Soler and Sordé, 2015).

The SIOR is peer-reviewed and projects are scored (1 to 10) according to the degree of fulfilment of a set of criteria: Connection to official social targets (such as United Nations Sustainable Development Goals or EU2020 targets); percentage of improvement achieved in relation to the starting situation; transferability of the impact; publication by scientific journals or by governmental or non-governmental official bodies; sustainability throughout time. Starting from these criteria, the research team is working on the development of more specific indicators. Projects already displayed in SIOR are providing valuable information and efforts for an accurate representation of their achievements.

Beyond gathering the evidences of impact, the SIOR is intended to contribute to the change of scientific research towards to the social improvement. Here, the feature of being Open Access is crucial, as it enables to promote open debate and reflection to the scientific community itself, but also to the society and policy makers. The first pilot of SIOR has consisted in a web application connected to a standard database. During the 2015 the research team has been testing this format and has also invited researchers from other disciplines and geographical areas to include their social impacts in the system. There has been developed the interoperability with ORCID, in the user registration and researcher profiles. In 2016, further developments will be done by using the free software for repositories DSPACE. One of the objectives is also to improve the interoperability in other research information systems.

These results are expected to contribute to place social impact of research in a greater status. On the other hand, the own process of development of the indicators, thorough the explained scientific activities and the permanent improvement of SIOR, is helping to transform its peripheral space.

REFERENCES AND CITATIONS

Flecha R., Soler M. & Sorde T. (2015). Metrics such as social improvement, transferability to diverse populations or social contexts, are also used to back research proposals. *Nature*, 528 (193). doi:10.1038/528193d



Je veux bien, mais me citerez-vous? On publication language strategies in an anglicized research landscape¹

Nadine Desrochers* and Vincent Larivière***

*nadine.desrochers@umontreal.ca; vincent.lariviere@umontreal.ca

École de bibliothéonomie et des sciences de l'information, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal 1, Québec, H3C 3J7 (Canada)

** Observatoire des Sciences et des Technologies (OST), Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Université du Québec à Montréal 1, CP 8888, Succ. Centre-Ville, Montréal, Québec, H3C 3P8 (Canada)

ABSTRACT

The pressure to publish in science's lingua franca is linked to a common belief that this choice will cause the research to be more readily indexed, accessed, read, used, and cited. However, the use of a national language can be marketed as a source of distinction for institutions located in countries or nations where English is not the primary language. This study looks to understand publication-language practices in the social sciences and humanities by examining the publication strategies of three nations, including a stateless nation: Germany, France, and Québec. The data were extracted from the Social Sciences Citation Index and the Arts & Humanities Citation Index and comprise 3.7 million articles, notes and reviews published between 1980 and 2014. The rise of English and decline of other languages is staggering and follows the same tendency in Germany and France, reaching just over 80% in each case. Québec differs slightly because the percentage of papers published in English was already quite high in 1980; nevertheless, the proportion has also risen, now reaching over 90%. Impact follows suit: for each of the three nations, papers published in English gather, on average, three times as many citations as they national-language counterparts. Given the reign of impact indicators and the symbolic capital granted to citations in the current scientific context, the data reveal that opting for English-centric publication strategies pays off. However, this raises questions fundamental to science, the symbolic capital associated with language, and the effects of language-based strategies on research.

¹ The authors thank the Agence universitaire de la Francophonie and the editors of *Découvrir* for making this research visible in French. This research-in-progress paper is part of an ongoing, broader study. Preliminary results have never been published in English; for French versions of previous steps, see: Desrochers, N. and Larivière, V. (In press). *Recognition ou reconnaissance : de la question des langues en diffusion des connaissances*. In Borg, S., Cheggour, M., Desrochers, N., Gajo, L., Larivière, V., & Vlad, M. (Eds.). *L'Université en contexte plurilingue dans la dynamique numérique: Actes du congrès annuel de l'Agence universitaire de la Francophonie, Marrakech, 12-13 novembre 2015*. Paris: Éditions des Archives Contemporaines; and a short column, "Langues et diffusion de la recherche: le cas des sciences humaines et sociales", in *Découvrir: le magazine de l'ACFAS*, <http://www.acfas.ca/publications/decouvrir/2015/11/langues-diffusion-recherche-cas-sciences-humaines-sociales>. For this paper, the choice was made to use Canadian English.

INTRODUCTION

In countries where the national or predominant spoken language is not English, “publish or perish” has a twist: “publish in English or perish”. However, for institutions located in such countries or nations, the use of the national language can be marketed as a distinction. For example, the Université de Montréal’s website states that it is “la seule université *francophone* canadienne à figurer parmi les 150 meilleurs établissements universitaires dans tous les classements internationaux” (Université de Montréal, n.d.; emphasis added). Furthermore, schemes established by funding agencies often prioritize local societal impact, in the social sciences in particular. However, the question of whether this translates into actual symbolic capital for researchers remains unanswered. This study looks to understand publication-language practices in the social sciences and humanities; for while a researcher’s national affiliation may be the result of a lifelong series of events, what to study, where to publish, and in what language are all, ultimately, choices.

BACKGROUND

One of the problematic concepts in studying the anglicization of science is that of the “internationalization” of publishing venues. For instance, Buela-Casal, Perakakis, Taylor and Checa (2006) concluded that, at least in psychology, « no single criterion provides an unequivocal measure of internationality » (p. 60). Other studies have looked at researchers’ perceptions of the migration towards English as the international language of science. Gnutzmann and Rabe’s (2014) qualitative analysis showed a mix of perceptions within the group of 24 German researchers interviewed. Schubert and Michels (2013) looked at the scientific impact of papers published by “large publisher nations” and found a parallel of the Mathew effect (Merton, 1968) for journals (see also Larivière and Gingras, 2010).

It is quite plain to see that the internationalization of the objects of study in the natural and medical sciences (NMS; Gingras and Mosbah-Natanson, 2010) has been accompanied by a gradual but undeniable migration towards English. The fact that these objects retain their intrinsic proprieties from one country to the next allowed these fields to fully embrace an internationalization of the dissemination of science (see Kirchik, Gingras and Larivière [2012] for a study of the effects of this in the Russian context). In fact, more than 98% of peer-reviewed documents in the NMS published in 2014 and indexed in WoS were in English.

On the other hand, objects of study for the social sciences and humanities (SSH) tend to have a more local focus (Warren, 2014). This leads us to ponder the symbolic capital (Bourdieu, 2001) that could—or perhaps that should—be given by institutions and policy makers to the dissemination of research in national languages.

Three nations, including a stateless nation, will be studied here in order to shed some light on the underlying tensions in publication strategies in the SSH: Germany, France, and Québec.

METHODS

The data were extracted from the Social Sciences Citation Index and the Arts & Humanities Citation Index of the WoS, which index 3,500 SSH journals. The dataset comprised 3.7

million articles, notes, and reviews (henceforth, “papers”) published between 1980 and 2014². Three variables were considered:

1. Country of affiliation of the first author;
2. Place of publication of the journal, established by the city provided in the address;
3. Language of the paper.

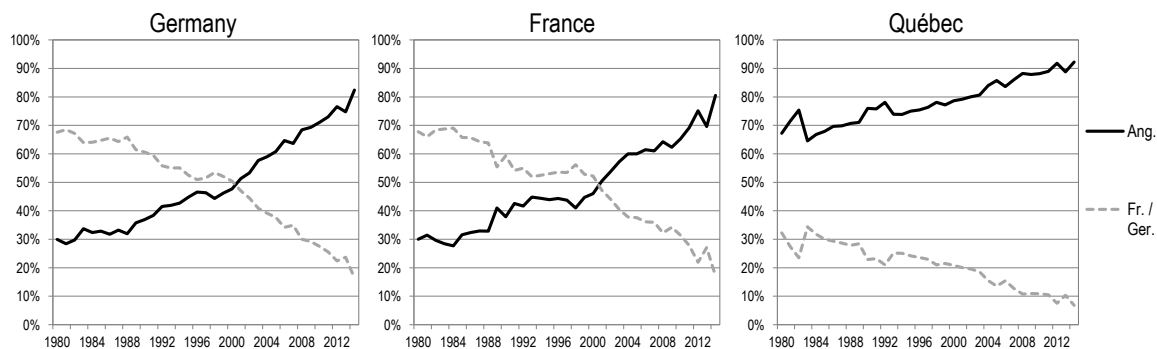
Scientific impact was obtained by using the number of citations received, normalized by year and by the journal’s discipline.

RESULTS

As shown in Figure 1, the rise of English and decline of national languages is staggering and follows the same tendency in Germany and France: currently, more than 80% of papers from these countries indexed in WoS are written in English, from roughly 30% 35 years ago. Less than 20% are being published in the national languages of German or French. Québec differs slightly because the percentage of papers published in English was already quite high in 1980; nevertheless, the proportion has also risen, now reaching over 90%. In absolute numbers, the dataset contains, for the most recent year (2014):

- For Germany: 8,644 papers in English and 1,718 papers in German;
- For France: 4,259 papers in English and 905 papers in French;
- For Québec: 1,986 papers in English and 147 papers in French.

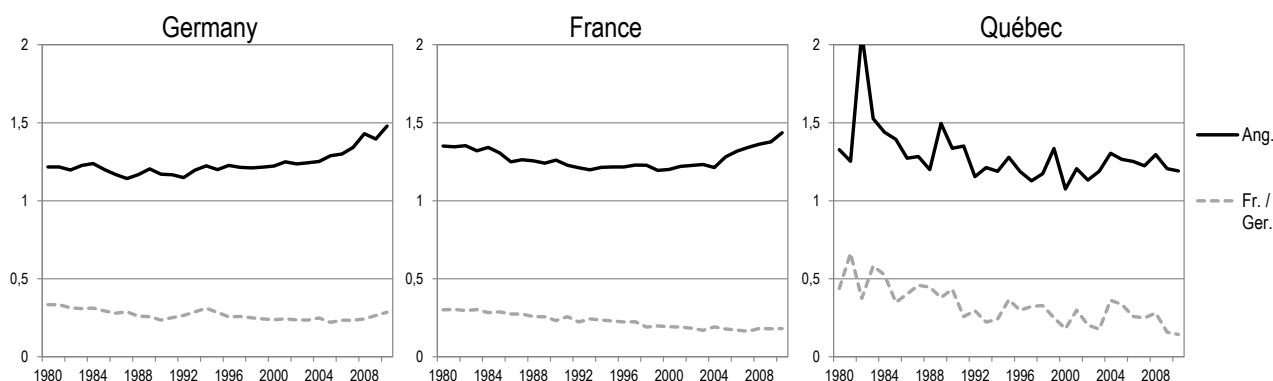
Figure 1. Percentage of papers in the social sciences and humanities written in English, German (for Germany) and French (for France and Québec), 1980-2014



Impact follows suit (Figure 2): for each of the three nations, papers published in English gather, on average, three times as many citations as they national-language counterparts; this tendency has been rising quite steadily since the turn of the millennium for Germany and France. The same can be seen for Québec, even though the fewer number of papers explains the wider variations.

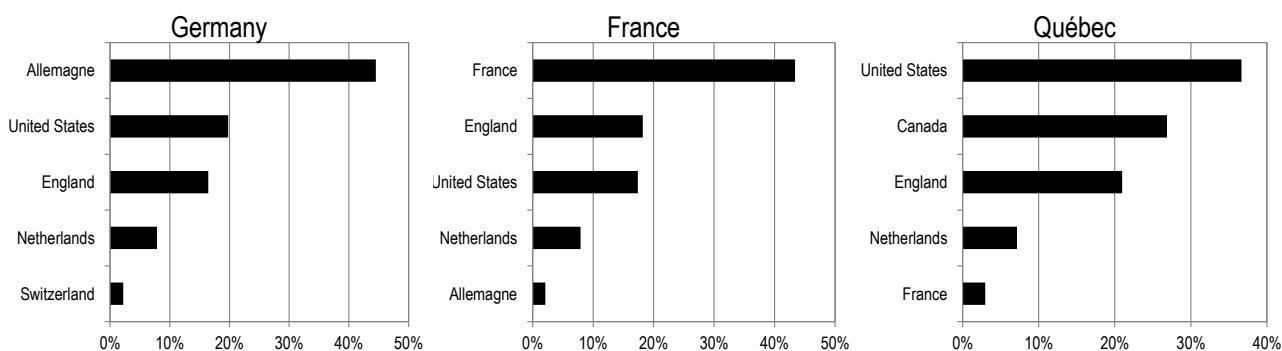
² WoS data for 2013 contains a high proportion of papers (roughly 50%) without publication language; these were excluded from the analysis.

Figure 2. Scientific impact of papers in the social sciences and humanities written in English, German (for German) and French (for France and Québec), 1980-2014



Interestingly, for Germany and France, publication venue choices are still predominately national, while American journals are the venues of choice for Québécois researchers (Figure 3). Also interesting is the fact that in terms of “large publisher nations” (Schubert and Michels, 2013), the positions of English and American journals are reversed for Germany and France, and that the Netherlands (home of Elsevier) comes in 4th place for all three nations. The fact that Québécois research is barely published in France also deserves to be noted.

Figure 3. Country of journals in which the German, French, and Québécois researchers publish in the social sciences and humanities, in percentages, 1980-2014



DISCUSSION AND CONCLUSION

The limitations of the study are inherent to the use of WoS data for the study of the social sciences and humanities (see Archambault et al., 2006; Larivière et al., 2006). These limitations notwithstanding, the clear tendency for German, French, and Québécois researchers to publish more and more in English is telling, as the data reveal that such a strategy pays off.

However, this raises questions fundamental to science: are there still contexts where opting for a language other than English can play in a researcher’s favour? What symbolic capital can be associated with publication in a national language? And how will this capital be measured in a researcher’s evaluation or in an institution’s ranking?

Language is more than a vehicle for knowledge; it has, in the words of Bourdieu, “symbolic power” (1991). In countries where language and identity are intertwined, language strategies become more than a question of Impact Factor; they are a question of impact in the broader sense of the word, of collisions and repercussions far-reaching both ideologically and practically. With all of its obvious advantages, the near-complete anglicization of science is nigh; and until policy makers, funding agencies, institutions, peer evaluators, and indicators align to grant clear value to national languages, researchers will likely continue to migrate to publish, in order not to perish.

REFERENCES

- Archambault, É., Vignola-Gagné, É., Côté, G., Larivière, V., & Gingras, Y. (2006). Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. *Scientometrics*, 68(3), 329-342.
- Bourdieu, P. (1991). *Language & symbolic power*. G. Raymond and M. Adamson, trans. Cambridge, MA: Harvard University Press.
- Bourdieu, P. (2001). *Science de la science et réflexivité*. Paris: Éditions Raisons d’agir.
- Buela-Casal, G., Perakakis, P., Taylor, M., & Checa, P. (2006). Measuring internationality: Reflections and perspectives on academic journals. *Scientometrics*, 67(1), 45-65. doi: 10.1556/Scient.67.2006.1.4
- Gingras, Y., & Mosbah-Natanson, S. (2010). Les sciences sociales françaises entre ancrage local et visibilité internationale. *European Journal of Sociology*, 51(2), 305-321.
- Gnutzmann, C. & Rabe, F. (2014). ‘Theoretical subtleties’ or ‘text modules’? German researchers’ language demands and attitudes across disciplinary cultures. *Journal of English for Academic Purposes*, 13(2014), 31-40. doi: 10.1016/j.jeap.2013.10.003
- Kirchik, O., Gingras, Y., & Larivière, V. (2012). Changes in publication languages and citation practices and their effect on the scientific impact of Russian science (1993–2010). *Journal of the American Society for Information Science and Technology*, 63(7), 1411-1419. doi: 10.1002/asi.22642
- Larivière, V., Archambault, É., Gingras, Y., & Vignola-Gagné, É. (2006). The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities. *Journal of the American Society for Information Science and Technology*, 57(8), 997-1004.
- Larivière, V. & Gingras, Y. (2010). The impact factor’s Matthew Effect: A natural experiment in bibliometrics. *Journal of the American Society for Information Science and Technology*, 61(2), 424-427. doi: 10.1002/asi.21232
- Merton, R.K. (1968). The Matthew effect in science—The reward and communication systems of science are considered. *Science*, 159(3810), 56-63. Retrieved from <http://www.jstor.org/stable/1723414>

Schubert, T. & Michels, C. (2013). Placing articles in the large publisher nations: Is there a “free lunch” in terms of higher impact? *Journal of the American Society for Information Science and Technology*, 64(3), 596-611. doi: 10.1002/asi.22759

Université de Montréal. (n. d.). Faits et chiffres. Retrieved from <http://recteur.umontreal.ca/documents-institutionnels/faits-et-chiffres/>



Effects of performance-based research funding on publication patterns in the social sciences and humanities¹

Raf Guns* and Tim C.E. Engels*

* raf.guns@uantwerpen.be, tim.engels@uantwerpen.be

University of Antwerp, Faculty of Social Sciences, Centre for R&D Monitoring (ECOOM), Middelheimlaan 1, 2020 Antwerp, Belgium

ABSTRACT

Publishing in the social sciences and humanities (SSH) and research evaluation practices are co-evolving. In this paper we present an analysis on how in Flanders the PRFS has shaped and influenced publication practices in the SSH. Our analysis is based on the VABB-SHW, a comprehensive database of research output in the SSH in Flanders.

We find that a strong emphasis on WoS publications since 2003 has caused a growth in WoS publications, that is greater than what can be observed in other countries and other fields of science in Flanders. Other mechanisms appear to exist for book publications, which are not indexed in the WoS databases used for the PRFS.

INTRODUCTION

In more and more countries and regions, government funding of universities and research institutions is linked to their research output (Hicks, 2012a). The effects of such performance-based research funding systems (PRFSs) have been studied in a number of papers.

Butler (2002, 2004) was among the first to show how researchers adapt their publication behaviour to changes in the system. Specifically, she studied the effects of changes in the Australian national funding model in 1993, which included productivity indicators that did not account for the quality of the publication or publication outlet. The main results showed that, while publication productivity increased, the overall citation impact of Australian publications decreased. Moreover, the relative increase was highest in lower-impact journals. Butler (2002) concluded that “the present system rewards quantity, not quality.”

The studies of Butler showed that, at least in some cases, researchers adapt their publication behavior to the parameters of the local funding model. Rousseau and Rousseau (2015) refer to this as metric-wisness, the phenomenon that researchers are aware of and can adapt to the indicators by which they are evaluated and/or funded.

Hicks (2012b) focusses on the position of the social sciences and humanities (SSH) in the context of performance-based research funding. In OECD countries, STEM fields account for 70–80% of government research spending. As a consequence, traditional publishing patterns in the SSH are under pressure, in order to better fit research evaluation protocols that are anchored mainly on the sciences. At the same time, in several cases it can be observed that there is a backlash against evaluation protocols that are too narrowly aimed at the sciences

¹ This investigation has been made possible by financial support of the Flemish government to ECOOM. We thank Jesper W. Schneider for sharing some of the data underlying Figure 1 and Linda Sile for useful comments on a previous version.

and the evaluation protocols are revised to better suit the characteristics of the SSH. In other words, research evaluation and SSH publishing are *co-evolving*.

A good example of the co-evolving dynamic described by Hicks (2012b) is the case of Flanders, the northern region of Belgium. In Flanders, the BOF-key is the distribution key used to determine how much funding each of the five Flemish universities receive out of the government University Research Fund (BOF), currently accounting for roughly 150 million euro. The BOF-key has included research output data from 2003 onwards (Debackere & Glänzel, 2004), based on publication and citation data from the Science Citation Index Expanded (SCIE), part of Web of Science (WoS). Because of the nature of the SCIE, research output of the SSH was almost invisible in the BOF-key. This spurred on two further reforms in 2008 (Spruyt & Engels, 2013; Verleysen et al., 2014):

- Four other WoS databases – the Social Sciences Citation Index (SSCI), the Arts and Humanities Citation Index (AHCI), as well as the conference proceedings databases – were included.
- The framework was laid out for constructing a comprehensive local database of research output in the SSH, the Flemish Academic Bibliographic Database for the Social Sciences and Humanities (VABB-SHW), which was first used in the BOF-key of 2011 (Engels, Ossenblok & Spruyt, 2012). Originally accounting for 2.6% of the BOF-key, the VABB-SHW currently represents 6.8% of the BOF-key.

These evolutions illustrate how SSH publishing has affected the PRFS. In this paper we will consider the reverse direction: effects of the PRFS on publishing patterns in SSH.

According to the BOF legislation, publications in the VABB-SHW should (1) be publicly accessible, (2) have an ISSN or ISBN, (3) contribute to the development of new insights or applications thereof, and (4) have been subjected to a demonstrable peer-review process by experts in the field. These criteria are upheld by the Authoritative Panel (Gezaghebbende Panel or GP), a panel of 18 professors affiliated to the five Flemish universities and coming from the different SSH disciplines (Verleysen, Ghesquière, & Engels, 2014). In addition to these four criteria, the GP has decided to exclude publications that count less than four pages.

Five publication types are included in the VABB-SHW. These types are weighted differently:

- journal articles: 1 point,
- books as author: 4 points,
- edited books: 1 point,
- book chapters: 1 point,
- proceedings papers: 0.5 points.

We can distinguish between two subsets of the VABB-SHW: *VABB-WoS* publications are those journal articles and proceedings papers that are also indexed in Web of Science (SCIE, SSCI, AHCI, CPCI-S, or CPCI-SSH), whereas *VABB-GP* publications are those publications that are not indexed in WoS but instead selected by the GP. Furthermore, we will also consider those publications that do not meet all the abovementioned criteria and are hence not approved for inclusion in the VABB-SHW. We will refer to this group as *Non-approved*.

In this paper we tentatively address the following questions:

1. How did changes in the PRFS in Flanders affect publication growth in WoS over all disciplines?
2. How did changes in the PRFS in Flanders affect publication growth in the SSH, both in WoS and in the VABB-SHW?

3. How did changes in the PRFS in Flanders affect evolutions in the relative share of different publication types?

Since it is very hard to establish causal links between changes in the PRFS and publication patterns, we consider the present analysis a first step to answering questions like the ones above.

DATA AND METHODS

Our main data source is the VABB-SHW, a comprehensive database of peer-reviewed publications from the SSH. We consider publications from the period 2000–2013 (inclusive). In Table 1 some basic descriptive statistics are provided about the data set used. The number of VABB-GP and VABB-WoS publications is similar for journal articles and proceedings papers, the two publication types where both groups occur. However, VABB-GP also includes book publications, which are not indexed in any of the WoS databases that are used for the BOF-key.

Table 1. Numbers of VABB-GP and VABB-WoS publications in period 2000–2013

	VABB-GP	VABB-WoS
Journal articles	23148	24681
Books as author	909	-
Edited books	1618	-
Book chapters	8953	-
Proceedings papers	1017	1158

To answer the first question, we also consulted data on total WoS publication output (including non-SSH publications) by Flemish universities, as used for the BOF-key.

Since we are interested in change and evolution over time, we will mostly work with relative rather than absolute numbers. More specifically, we take the first year of a time period as a point of comparison and then observe how different sets evolve relative to their value in the first year.

RESULTS

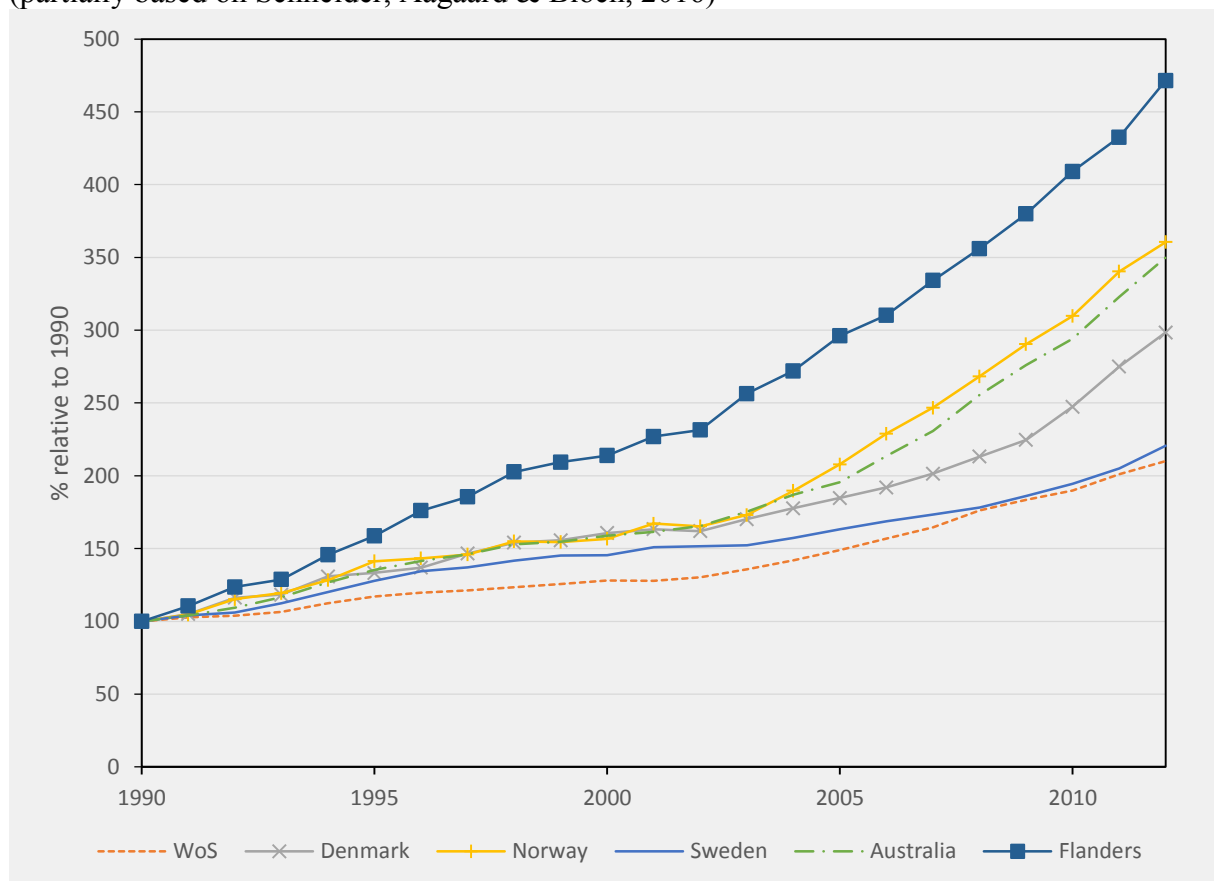
The BOF-key has included publications (and citations) in WoS as a parameter since 2003. Figure 1 compares the evolution of the number of WoS publications in Flanders with that in four other countries over the course of a 23-year period. We observe that in all countries the increase in publications is greater than the growth of WoS itself (the lowest line). The growth in Flanders is considerably stronger than in the other countries. From 2003 onwards, the increase becomes even steeper. This suggests that the introduction of a parameter of research output that is entirely based on the WoS, has driven researchers to publish more in WoS-indexed publication outlets. At different points in time similar trend changes can be observed for Norway, Australia, and Denmark.

Between 2000 and 2012, the amount of WoS publications by researchers affiliated to Flemish universities has grown by 230%. This increase cannot be ascribed solely to the PRFS. Other factors are at play, including the following:

- The WoS itself has increased its coverage of journals and proceedings.
- The number of Flemish researchers has increased by 175% over this same period. Especially the group of pre- and postdoctoral researchers has grown, while the number of professors has expanded only slightly.
- Flanders is no exception to the global trend toward more collaboration (Ossenblok, Verleysen & Engels, 2014). This is an important factor that helps to explain increased publication volume (Fanelli & Larivière, 2016).
- Better monitoring of publication output by individuals, research groups, institutions and the government alike

Of course, these factors also occur to some extent in other regions and countries.

Figure 1. Evolution of WoS publications relative to 1990 in different countries (partially based on Schneider, Aagaard & Bloch, 2016)

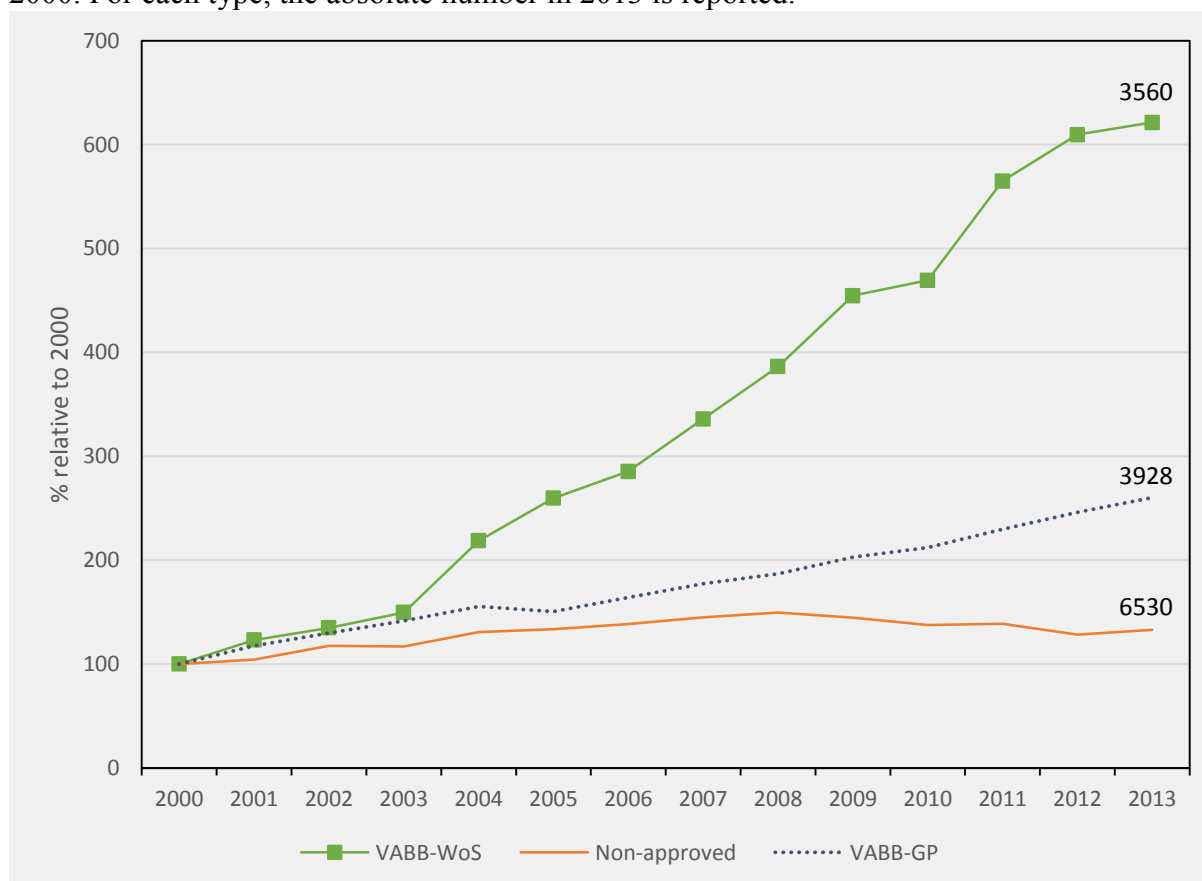


We now focus on evolutions that can be observed within the SSH, using data from the VABB-SHW. As one might expect, we witness an increase (Figure 2) between 2000 and 2003 for all three sets of publications – VABB-WoS, VABB-GP, and Non-approved. Between 2000 and 2003, the evolution of peer-reviewed literature within (VABB-WoS) and outside of WoS (VABB-GP) appears to run in parallel, but the number of WoS publications grows faster than that of GP publications afterwards. The strong rise of WoS publications observed for Flanders across all disciplines can also be seen within the SSH, with a „jump“ starting in 2003. In fact, the growth of WoS publications is even stronger within the SSH. Factors that may

help to explain this finding include: the introduction of the PRFS in 2003; the expansion of the PRFS with the AHCI and the SSCI in 2008; the fact that several Flemish and Dutch journals were added to the WoS in the period 2005–2009 (Ossenblok, Engels, & Sivertsen, 2012) as part of Thomson Reuters' coverage expansion to regional literature (Testa, 2011); and, more generally, the low volume of WoS publications in 2000 ($n=573$ or 33% of all journal articles and proceedings).

We further observe that the growth of VABB-GP publications is faster than that of Non-approved publications, both before and after the introduction of the VABB-SHW. Around 2008 the curve of Non-approved publications reaches a maximum and slightly decreases in the years thereafter. In 2008 work also began on the construction of the VABB-SHW, although it seems unlikely that this could have immediately brought about this change. In spite of the different evolution of peer-reviewed and non-peer-reviewed literature, we note that in absolute terms the number of Non-approved publications in 2013 is still slightly higher than the number of VABB-WoS and VABB-GP combined.

Figure 2. Evolution of VABB-WoS, VABB-GP and non-approved publications relative to 2000. For each type, the absolute number in 2013 is reported.



The increasing trend for VABB-GP publications one can observe in Figure 2 is not the same across the five different publication types. Figure 3 shows the evolution for the five publication types (restricted to VABB-GP) and adds the curve for WoS articles for comparison purposes. Because the volume of proceedings is very low, GP and WoS proceedings are taken together.

Let us first look at journal articles. The difference between journal publications that are and are not indexed in WoS is striking. Researchers are increasingly opting for journal publications in WoS, whereas the growth for other peer-reviewed journals is much slower. This is not entirely unexpected, given the fact that until 2011 only WoS-indexed articles were taken into account for the BOF-key. However, the introduction of the VABB-SHW and of GP-indexed journal articles as (part of) a parameter in the BOF-key does not appear to have caused a divergence of the existing trend.

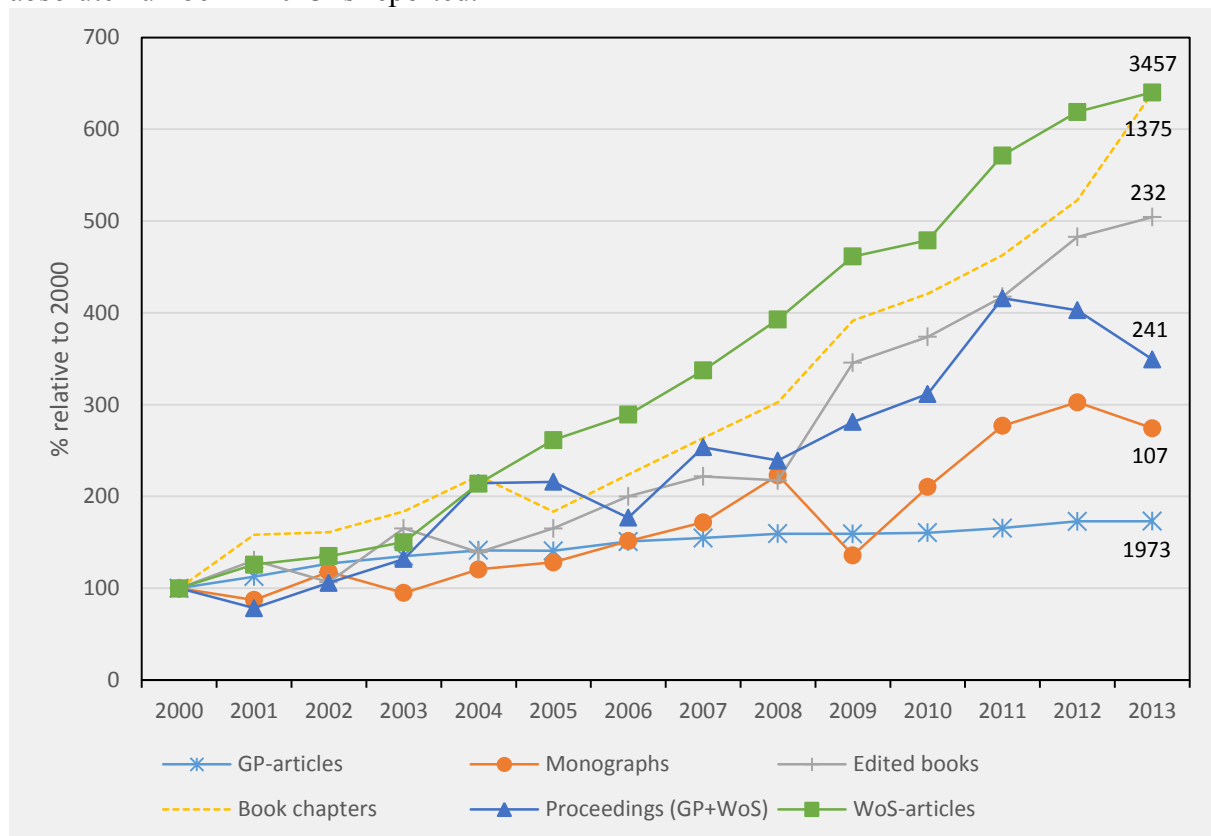
Table 2. Number of WoS and GP journals in which SSH scholars working in Flanders have published

	WoS journals	GP journals
2000–2002	1077	1127
2006–2008	2165	1715
2011–2013	3125	1944

A similar picture emerges if we consider journals rather than journal articles. Table 2 shows the evolution of number of WoS and GP journals in three three-year time periods. The (non-cumulative) number of WoS journals in which SSH scholars working in Flanders have published has almost tripled between the first and the last time period, whereas the number of GP journals has grown by a factor of 1.7.

It is quite striking that edited books and especially book chapters in Figure 3 follow almost the same trend as VABB-WoS articles. This could be interpreted as the PRFS having little to no influence on publication type. On the other hand, we could also see it as a result of the fact that some publication types are partially indexed in WoS, whereas others are not at all. Indeed, for journal articles the difference between VABB-GP and VABB-WoS is very clear. For those publication types to which the criterion „indexation in WoS“ does not apply, the criteria used by the GP are of greater relevance. The increase in book publications is likely also related to the introduction of the GPRC label for peer-reviewed books in 2010 (Verleysen & Engels, 2013).

Figure 3. Evolution of publication types in VABB-SHW relative to 2000. For each type, the absolute number in 2013 is reported.



CONCLUSIONS

In this paper we present an analysis on how in Flanders the PRFS has shaped and influenced publication practices in the SSH. We find that a strong emphasis on WoS publications since 2003 has caused a growth in WoS publications that is greater than what can be observed in other countries and other fields of science in Flanders. The introduction of the VABB-SHW has not led to a marked decrease of WoS publications in the SSH. Instead, WoS articles have grown 3.5 times faster than GP-selected articles over the course of the 14-year time period considered in this paper.

Other mechanisms appear to exist for publication types that are not indexed in the WoS databases used for the PRFS. This is the case for edited books and especially book chapters, which witness a growth comparable to that of WoS articles.

REFERENCES

- Butler, L. (2002). A list of published papers is no measure of value. *Nature*, 419(6910), 877.
- Butler, L. (2004). What happens when funding is linked to publication counts? In: Moed, H. et al. (eds.), *Handbook of Quantitative Science and Technology Research*, pp. 389–405. Netherlands: Kluwer.
- Debackere, K. & Glänzel, W. (2004). Using a bibliometric approach to support research policy making: The case of the Flemish BOF-key. *Scientometrics*, 59(2), 253–276.

- Fanelli, D., & Larivière, V. (2016). Researchers' individual publication rate has not increased in a century. *PloS One*, 11(3), e0149504. <http://doi.org/10.1371/journal.pone.0149504>
- Hicks, D. (2012a). Performance-based university research funding systems. *Research Policy*, 41(2), 251–261.
- Hicks, D. (2012b). One size doesn't fit all: On the co-evolution of national evaluation systems and social science publishing. *Confero: Essays on Education, Philosophy and Politics*, 1, 67–90.
- Ossenblok, T. L. B., Engels, T. C. E., & Sivertsen, G. (2012). The representation of the social sciences and humanities in the Web of Science - A comparison of publication patterns and incentive structures in Flanders and Norway (2005-9). *Research Evaluation*, 21(4), 280–290.
- Ossenblok, T. L. B., Verleysen, F. T., & Engels, T. C. E. (2014). Coauthorship of journal articles and book chapters in the social sciences and humanities (2000-2010). *Journal of the Association for Information Science and Technology*, 65(5), 882–897.
- Rousseau, S., & Rousseau, R. (2015). Metric-wiseness. *Journal of the Association for Information Science and Technology*, 66(11), 2389.
- Schneider, J.W., Aagaard, K., & Bloch, C.W. (2016). What happens when national research funding is linked to differentiated publication counts? A comparison of the Australian and Norwegian publication-based funding models. *Research Evaluation*, in print. doi:10.1093/reseval/rvv035.
- Spruyt, E.H.J., & Engels, T.C.E. (2013). Nieuwe sleutel verdeling van middelen Bijzonder Onderzoeksfonds. *Thema: Tijdschrift voor hoger onderwijs en management*, 13(3), 56–61.
- Testa, J. (2011). *The globalization of Web of Science: 2005-2010*. Philadelphia: Thomson Reuters. http://wokinfo.com/products_tools/multidisciplinary/webofscience/contentexp/expansionessay/
- Verleysen, F. T., & Engels, T. C. (2013). A label for peer-reviewed books. *Journal of the American Society for Information Science and Technology*, 64(2), 428-430.
- Verleysen, F. T., Ghesquière, P., & Engels, T. C. E. (2014). The objectives, design and selection process of the Flemish Academic Bibliographic Database for the Social Sciences and Humanities (VABB-SHW). In W. Blockmans, L. Engwall, & D. Weaire (Eds.), *Bibliometrics. Use and Abuse in the Review of Research Performance* (pp. 117–127). Portland Press. <http://www.portlandpress.com/pp/books/online/wg87/087/0117/0870117.pdf>
-



Developing appropriate methods and indicators for evaluation of research in the social sciences and humanities.

PRESENTATION OF A NEW COST ACTION

As STI ENID 2016 will focus on areas of research which are traditionally treated as "peripheral" in bibliometrics because they are inadequately covered or targeted by current international data sources and indicators, one of the major issues that comes to mind is that of SSH research evaluation. Based on a newly accepted COST Action, we propose to organize a roundtable devoted to the analysis of specific problems related to indicator use in SSH research evaluation, and of new, creative uses of metrics for this area.

CONTEXT

The humanities, at the origin of European universities (law, philosophy, religious studies, language and history), continue to attract large numbers of students, together with the social sciences, deliver highly educated people, produce world class research and contribute significantly to the financial viability of the European university system. At the same time, they struggle to adapt to impact evaluation systems that do not fully reflect the aspirations and research patterns of SSH disciplines because these tend to emphasize impact in terms of immediate economic returns, rather than focus on other societal returns from these domains of academic scholarship. Moreover, since the nineties of the previous century (Nederhof, 1989), studies have been developed showing that SSH output is, for a very large part, invisible through the large international databases which focus on scientific publications (Web of Sciences and Scopus), and which are often used for metric informed decision making in research programs and policies (Haddow & Genoni, 2010; Sivertsen & Larsen, 2012; Sivertsen & van Leeuwen, 2012; Sivertsen, 2016). Consequences of this are numerous, as most SSH research output is for example not taken into account amongst the criteria of the major league tables for the universities, like the ARWU (Shanghai) ranking or the Leiden Ranking. The effects of this lower visibility are multiple, spanning from a lower consideration and appreciation in academia to closing of SSH departments and drastic personnel reduction in certain universities (e.g., in Japan or the USA)¹. Aware of this problem, a number of countries have adopted qualitative assessment procedures for research evaluation in their assessment protocols, but the strong correlation between the funding discourse and quantitative performance indicators and measures remains a reality, preventing the development of SSH full potential. This is detrimental for the SSH fields, but also for societies that are challenged by a number of problems that are in need of SSH knowledge (migration crises, religious conflicts, economic failures, political system crises).

Particularly problematic for SSH domains are the effects of the failure to demonstrate economic returns, at a time when economic and societal impact is becoming an increasingly pervasive element of research funding argument, resulting in reduced research funding made available for the SSH. Much knowledge exists about SSH impacts, but, because of their lack of concreteness (maintaining and improving the resilience, coherence and adaptive capacity of societies) and heterogeneity, it has proven difficult to upscale this value and to demonstrate it clearly, in the way that other disciplines have been able to do through a (admittedly partial)

¹ <http://www.nytimes.com/2016/02/22/business/a-rising-call-to-promote-stem-education-and-cut-liberal-arts-funding.html?ref=topics&r=0>

set of metrics, counting spin-offs, licenses and patents as evidences of economic and societal impact.

ENRESSH: A COST ACTION TO JOIN FORCES

The challenge is therefore to enable the SSH to better demonstrate their true place in academia and society. To do so, scholars from different fields and backgrounds with a strong research record in the field of research on the SSH have joined forces in a new COST Action - ENRESSH: European network for Research Evaluation in the Social Sciences and Humanities. The Action proposes to bring together different strands of work consecrated to SSH research evaluation, currently under development in different parts of Europe, in order to gain momentum, to exchange best practices and results, and to avoid unnecessary duplication. Its main aims, deeply interrelated, are to:

- 1) improve evaluation procedures in order to take into account the diversity and the wealth of SSH research;
- 2) make a robust case for the ways in which the SSH add value to the society;
- 3) help SSH scholars better appropriate their research agenda and overcome fragmentation.
- 4) open up SSH research for interaction with societal stakeholders.

These outcomes are essential if Europe is to fully benefit from one of its historical and major assets in the competition to become the world leader of the knowledge economy (society) and to attract international students. They are also a key for further building the European Research Area. The Action has two main objectives, the first consists of the progress in the design of research evaluation procedures specifically for SSH fields, including the assessment of their societal relevance and impact. Therefore, it is necessary to improve, above all, the understanding of how SSH fields generate knowledge, what kind of scientific and societal interactions characterize different SSH disciplines, and what are the patterns of dissemination in the SSH (*“the research coordination objective”*). The second objective focuses on a further integration of researchers of various backgrounds, whose tools and methods can help tackling the complex problems of SSH evaluation. The Action will therefore act as a platform for putting together research teams able to act as pool of specialists upon whom external stakeholders (evaluation agencies, policy makers, HEI’s managers and directors) may call to solve questions linked to the evaluation of SSH research (*“the capacity building objective”*). The contribution at the 2016 STI ENID Conference in Valencia will be to further strengthen the network of researchers working on these topics.

The development of both these objectives can profit from knowledge and expertise not only from researchers, but also from stakeholders who have to work with evaluation systems in practice. In that sense, the aim of the Action is to co-create more adequate systems for the humanities and social sciences.

ORGANIZATION OF THE ROUNDTABLE

The roundtable will propose four presentations, the first one closely related to the topic of the conference, while the three following presentations from the perspectives of the Working Groups will try to answer to the issues raised.

- 1) the first one summarizes examples of the use and abuse of metrics in SSH evaluation (use and abuse of metrics in assessing scholars activity; use and abuse of metrics in assessing excellence in the SSH; use and abuse of metrics in assessing societal impact, shortages in data collection supporting research assessment in SSH and law).
- 2) **WG 1. Conceptual frameworks for SSH research evaluation:** The objective of this working group is to further develop our understanding of the SSH knowledge production processes and strategies, as a basis for developing evaluation procedures

that adequately reflect the research practices, goals and aims of the SSH scholars. The working group will tackle the dialectic issues of the potentials and drawbacks of (a) metric approaches and peer review; (b) international exchange and cooperation and the local rootedness of SSH; and (c) the need for interdisciplinary exchange and disciplinary expertise.

- 3) **WG 2. Societal impact and relevance of the SSH research:** The objective of this working group is to analyze the non-academic partnerships and environments of SSH research, in their diversity.
- 4) **WG 3. Databases and uses of data for understanding SSH research:** The main objective of this working group is to reflect upon the standardization and the interoperability of current research information systems dedicated to the SSH research outcomes.

The roundtable at the conference aims to give an overview on different strands of work going on in this field of research and to bring together different stakeholders from research evaluation attending the conference, thus obtaining feedback from both international experts in the field of bibliometrics, as well as from other stakeholders.

REFERENCES

- Haddow, G., & Genoni, P. (2010). Citation analysis and peer ranking of Australian social science journals. *Scientometrics*, 85(2), 471-487.
- Nederhof, A.J., (1989). The validity and reliability of evaluation of scholarly performance. In: *Handbook of Quantitative Studies of Science and Technology* Edited by A.F.J. van Raan, Elsevier Science Publishers.
- Sivertsen, G., & Larsen, B. (2012). Comprehensive bibliographic coverage of the social sciences and humanities in a citation index: An empirical analysis of the potential. *Scientometrics*, 91(2), 567-575.
- Sivertsen, G. & van Leeuwen, T. (2014). *Science, Technology & Innovation Indicators 2014. Thematic Paper 2. Scholarly publication patterns in the social sciences and humanities and their relationship with research assessment*. Utrecht: Dialogic.
- Sivertsen, G. (2016). Patterns of internationalization and criteria for research assessment in the social sciences and humanities. *Scientometrics*. DOI: 10.1007/s11192-016-1845-1



Clashing Conventions? Exploring Human Resource Management in the Cleavage Between Academic Field Traditions and New Institutional Rules. Quantitative and Qualitative Insights from the Field of Communication and Media Studies in Switzerland¹

Alexander Buhmann^{*}, Benedetto Lepori^{**} and Diana Ingenhoff^{***}

^{*}alexander.buhmann@bi.no

Department of Communication and Culture, BI Norwegian Business School, Nydalsveien 37, Oslo, 0484 (Norway)

^{**}benedetto.lepori@usi.ch

Centre of Organizational Research, University of Lugano, Via Lambertenghi 10A, Lugano, 6904 (Switzerland)

^{***}diana.ingenhoff@unifr.ch

Department of Communication and Media Research, University of Fribourg, Bd de Perolles 90, Fribourg, 1700 (Switzerland)

ABSTRACT

Contemporary developments in the academic profession, such as increasing financial constraints, processes of differentiation, marketization, and rising international co-operation and competition, have a bearing on human resource management in higher education and research institutions (Enders & Musselin, 2008; Enders & Teichler, 1997). One aspect of these trends is the introduction of university-level rules on hiring and career development, which largely follow new public management approaches and focus on competition, mobility and performance as key criteria for hiring and internal promotions (Kehm & Teichler, 2013; Goastellec, Park, Ates & Toffel, 2013). At the same time universities remain largely decentralized organizations and much of the personnel selection takes place at the level of departments (Bleiklie, Enders & Lepori, 2015).

This can lead to conflicting pressures because new university standards may cut across various faculties and departments, replacing established conventions of academic fields. In most cases, the actors who make the institutional rules about hiring and career development are not the ones making the human resources decisions as recruitment of new colleagues and career development remains under the control of the heads of individual academic units (Musselin, 2005).

In this paper we want to study human resource management in the academic profession by focusing on the tension between institutional rules and academic field traditions in the case of a typical field in human and social sciences in Switzerland, i.e., communication and media studies (CMS). To this aim we combine three types of information: descriptive information on rules and procedures for hiring and career development at the institutional level, quantitative data on career paths⁷ in the field and, finally, qualitative interviews with heads of academic units from the field. We focus specifically on the analysis of the prevalent career model and on exploring possible tensions between university strategies on the one hand and the process of hiring and career development controlled by heads of academic units on the other hand. We finally discuss the implications of our findings for career management in universities.

¹ This work was supported by the Rectors' Conference of the Swiss Universities (CRUS)

LITERATURE REVIEW

As in most countries, universities in Switzerland are increasingly transformed from an earlier form of collegial communities of academics into hierarchical organizations (Boer & Goedegebuure, 2001; Brunsson & Sahlin-Andersson, 2000). At the same time, Swiss universities retained some essential elements of the academic governance, such as a high level of autonomy of academics and the decentralization of decisions concerning teaching and research (Bleiklie, Enders, Lepori & Musselin 2011).

An important dimension of this transformation has been the formalization of hiring procedures and human resources management, particularly in the leading research universities (Horta, 2009). Job roles and work tasks are becoming more differentiated and aligned on those of the business sector, with university leaders adopting the roles of managers (Rhoades & Sporn, 2002). New models of management introduce overarching university strategies and rules. These rules concern not only the specifics of career paths' but also the general relation between academics and their university. Both may affect recruitment and career development practices of the heads of academic units. As universities seek more flexible forms of employment, new standards on career paths' tend to create new positions around the more traditional academic career ladders (Enders & Musselin, 2008).

Due to the cultural diversity in Switzerland at the intersection between the German-, French-, and Italian-speaking region, two career path models were common before reforms: the survivor model (typical also of the Humboltian chair system in Germany or Austria) where PhDs must go through various trials to provide evidence of their talents and wait many years to obtain a permanent position; and the protective pyramid (typical also of the public systems in France or Spain) where PhDs gain access to different hierarchical categories of permanent positions quite early after a highly selective tournament and with the option to be promoted towards the top of the hierarchy depending on the growth rate of the overall pyramid and the age/seniority of those on the top.

In the last two decades, most Swiss universities, however, increasingly introduced regulations favoring the tenure model (typical of the liberal systems, e.g., in the US or UK), based on early selection of young PhDs among whom some are offered time-limited posts leading, at the end of a certain period of time, to a tenure procedure to decide whether they will be offered a tenured position. In addition to these changes in standards on career paths' there is a push towards an employer-like mode of universities in which academics tend to become "managed professionals" (Rhoades & Slaughter, 1997; Slaughter, 1997). These developments, however, bring some contradictions as the new rules aim not only at reinforcing academics' affiliation to the institution but simultaneously promote mobility and flexibility (Enders & Musselin, 2008).

Research Questions

As we aim to analyze the prevalent career paths' in Swiss CMS and explore possible tensions between university strategies and human resource management of heads of academic units, we focus on the following questions:

- How are careers organized?
- What is the basic career model behind hiring and career development strategies?
- Do career paths comply with the conventions in the field?
- What is the impact of university regulations on the hiring decisions and career development strategies of unit heads?

METHOD

We combine descriptive statistical evidence on career paths in Swiss CMS with qualitative interviews with unit heads on their hiring and career development strategies. First, we collected data on individual researchers in Swiss CMS for a five-year period (2009–2013). Websites, CVs and self-maintained web portraits of researchers were used to populate a list of individuals recording earned degrees, current and previous institutional affiliation (with start and end date), full-time equivalent positions, function, internal mobility and exit mobility. All data was sent to the heads of the institutes for validation. Second we conducted semi-structured interviews with the unit heads in charge of hiring and career development. A total of 17 out of 21 unit heads participated. Interviews were conducted between October 2015 and January 2016.

PRELIMINARY RESULTS

We combine descriptive statistical evidence on career paths' in Swiss CMS with qualitative interviews with unit heads on their hiring and career development strategies. First, we collected data on individual researchers in Swiss CMS for a five-year period (2009–2013). Websites, CVs and self-maintained web portraits of researchers were used to populate a list of individuals recording earned degrees, current and previous institutional affiliation (with start

University Regulations

The decentralized structure of Swiss higher education, where universities are largely free to decide on hiring procedures and human resource management, implies that there is no national framework for academic careers and that there are differences between individual institutions. However, some general trends exist. They include deregulation of the hiring of research staff at the post-doctoral and PhD levels based on temporary contracts decided by subunits (subject to availability of resources); an increasing formalization of the hiring of professor staff, where general rules are formulated at the university level but the hiring process takes place at the departmental/faculty level; the introduction of formal rules for careers, such as the limitation of hiring period for non-professor staff; the increasing importance of international mobility as hiring criterion; and the widespread introduction of a tenure model, where hiring takes place at the assistant professor level.

Quantitative Insights

Preliminary analyses of data on personnel structure and mobility patterns shows two different layers, each with distinct sources and hiring practices (c.f., Figure 1, with full-time positions, FTE, listed as five-year averages). The first layer at the early and intermediate level is characterized mostly by internal careers. Individuals tend to do their MA, PhD, and post-doc within the national field and in many cases within the same unit. This layer is complemented by external hires at the post-doctoral level.

The second layer at the professor level is mostly composed of external hires with 84% of professors entering from the outside. The early and intermediate researchers' layer is therefore much less internationalized than the professor layer. Exits from the first level occur at two stages: the first during or after the PhD for those leaving academia (38% of all PhD exits) and the second stage after a post-doc period for those who remain in academia. 68% accepted academic promotions outside the perimeter.

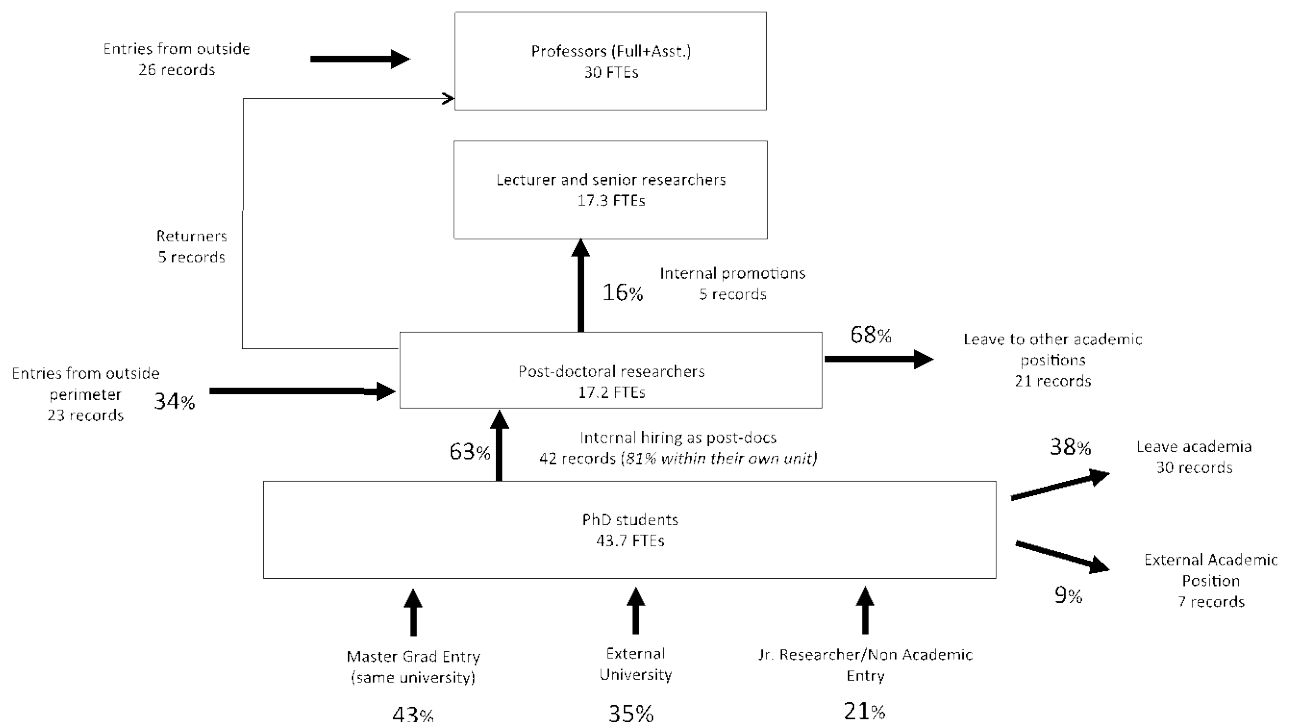


Figure 1. Career paths' in Swiss CMS

The striking feature of the career structure is therefore the clear divide between the two layers: in the first layer, internal reproduction prevails, while in the professor level most hires are from outside the country (and the field). The few people who managed to pass from the first to the second layer had a research experience or an appointment abroad.

Qualitative Insights

Preliminary analyses of the interviews reveal that the heads of academic units, to large extent, do not perceive an influence of the overarching university strategy on their strategies of developing the academic unit. On aggregate level universities overall strategy appear to be largely irrelevant for the field in terms of hiring and career development. Additionally, unit heads' strategies for the first layer (PhD and post-doc) are very heterogeneous ranging from clear emphasis on internationalization, co-operation, and marketization to a purely local or regional focus.

DISCUSSION AND CONCLUSION

Though the traditional consensus within universities about what it means to develop the academic profession is disappearing and new institutional rules cut across filed traditions on hiring and career development, our study shows that, in CMS, the model of academic careers which is introduced at the institutional level, is largely disregarded by academics when hiring researchers at the PhD and post-doc level: internal hires and inbreeding are frequent, while international mobility is given a low priority. Institutional rules are however much more forceful in the hiring of professors since regulations tend to be much more detailed and the university central level has a direct control power on nominations. The outcome of this (power and cultural) clash is that the researchers' layer in the field is largely cut-off from the

upper layer of university careers. Specifically, our study shows that Swiss CMS does not reproduce in its national context but relies heavily on external hires to the professor level.

REFERENCES

- Bleiklie, I., Enders, J. & Lepori, B. (2015). Organizations as penetrated hierarchies. *Organization Studies*, 36(7), 873–896.
- Bleiklie, I., Enders, J., Lepori, B. & Musselin, C. (2011). New public management, network governance and the university as changing professional organization. In T. Christensen & P. Lægred (Eds.) *Ashgate Research Companion to New Public Management*. Aldershot: Ashgate.
- Boer, H. D., & Goedegebuure, L. (2001). On Limitations and Consequences of Change: Dutch University Governance in Transition. *Tertiary Education and Management*, 7(2), 163–180.
- Brunsson, N., & Sahlin-Andersson, K. (2000). Constructing Organizations: The Example of Public Sector Reform. *Organization Studies*, 21(4), 721–746.
- Enders, J., & Musselin, C. (2008). Back to the future? The academic professions in the 21st century. In OECD (Ed.), *Higher education to 2030. Volume 1: Demography* (pp. 125–150).
- Enders, J., & Teichler, U. (1997). A victim of their own success? Employment and working conditions of academic staff in comparative perspective. *Higher Education*, 34(3), 347–372.
- Finkelstein, M. (2003, October 2). The Morphing of the American Academic Profession. Retrieved March 10, 2016, from <https://www.aacu.org/publications-research/periodicals/morphing-american-academic-profession>
- Goastellec, G., Park, E., Ates, G. & Toffel, K. (2013). Academic markets, academic careers: Where do we stand? In B. M. Kehm, Barbara & U. Teichler (Eds.), *The Academic Profession in Europe: New Tasks and New Challenges* (pp. 93–120). Dordrecht: Springer.
- Horta, H. (2009). Global and national prominent universities: internationalization, competitiveness and the role of the State. *Higher Education*, 58(3), 387–405.
- Kehm, B. & Teichler, U. (2013). *The Academic Profession in Europe: New Tasks and New Challenges*. Dordrecht: Springer.
- Musselin, C. (2005). *Le marché des universitaires : France, Allemagne, Etats-Unis*. Paris: Les Presses de Sciences Po.
- Rhoades, G., & Slaughter, S. (1997). Academic Capitalism, Managed Professionals, and Supply-Side, Higher Education. *Social Text*, (51), 9–38.
- Rhoades, G., & Sporn, B. (2002). New models of management and shifting modes and costs of production: Europe and the United States. *Tertiary Education and Management*, 8(1), 3–28.

Slaughter, S. (1997). *Academic Capitalism: Politics, Policies, and the Entrepreneurial University*. Baltimore: The Johns Hopkins University Press.



A bibliometric indicator with a balanced representation of all fields

Gunnar Sivertsen

gunnar.sivertsen@nifu.no

Nordic Institute for Studies in Innovation, Research and Education (NIFU)
P.O. Box 2815 Tøyen, N-0608 Oslo, Norway

ABSTRACT

As research in progress, we present two studies aimed at redesigning the bibliometric indicator of the “Norwegian Model” as response to an evaluation in 2013. The indicator is supposed to give a balanced representation of all fields, also those that are constructed as “peripheral” in traditional bibliometrics because of limited coverage in databases. The first study deals with balancing between different field-dependent co-authorship practices in the indicator, the other with the possible addition of a measurement of citation impact that could be applicable across all fields.

Keywords

Bibliometric indicators; productivity; citation impact; co-authorship; fractionalization; publication patterns; evaluation; the Norwegian model.

Submission type: Research in progress paper.

Relevant track: Social sciences and the humanities.

INTRODUCTION

The so-called “Norwegian Model” (Schneider 2009; Sivertsen 2010; Ahlgren et al. 2012), which so far has been adopted at the national level by Denmark, Finland and Norway, partly also by Belgium (Flanders) and Portugal, and at the local level by several Swedish universities, includes a bibliometric indicator which is intended to represent all fields of research comprehensively and comparably. It covers peer-reviewed scholarly and scientific publishing in books as well as in journals and series, and in national as well as international languages. Fields of research that are traditionally treated as “peripheral” in bibliometrics, such as the social sciences and the humanities, are thereby intended to be covered just as adequately as the sciences (Sivertsen 2016a).

After “light” assessments of the experiences with the model in the Danish and Flemish contexts in 2012 (Sivertsen & Schneider 2012; Technopolis Group 2013), the Norwegian model was evaluated extensively in Norway in 2013 with regard to its design, effects, organization, and legitimacy (Aagaard et al. 2014). As well as advising improvement and further development, the exercise provided the basis for four in-depth studies of internationally relevant questions (Aagaard 2015; Aagaard et al 2015; Schneider et al. 2015; Bloch & Schneider 2016).

This paper presents two studies with the aim to redesign the indicator in response to the evaluation. The first study deals with balancing between different field-dependent co-authorship practices in the indicator, the other with the possible addition of a measurement of citation impact in the indicator. The Norwegian government has already implemented the results of the first study as of 2016. The second study was recently commissioned by the government and will be finished before summer with results ready to be presented at the STI conference in September.

STUDY A: BALANCING THE MEASUREMENT OF PRODUCTIVITY ACROSS ALL FIELDS

The evaluation (Aagaard et al. 2014) found an imbalance in the indicator's representation of productivity across fields, thereby confirming an observation in an earlier study (Piro et al. 2013). The humanities and social sciences seemed to be more productive than the sciences, probably because the indicator uses fractional counting of co-authored publications. Even if co-authorship practices seemed unaffected at the macro level, the evaluation also expressed concern about the risk of discouraging research collaboration when using fractional counts (Bloch & Schneider 2016).

The indicator needs to be balanced because it is used for measuring productivity across institutions with different research profiles (e.g. general versus technical universities, universities with and without medical faculties). To study the problem and simulate its solution, we used data from CRISTin (Current Research Information System in Norway), covering the two years 2011-2012. In this system, authors are identifiable as real persons at Norwegian institutions, not just as author names with addresses. We studied the average productivity among 14,441 active researchers who had contributed to a minimum of two publications in the two years. The results are presented in Table 1 and explained below.

Table 1. Average number of publications (full counts) and publication points (old and new) per active researcher in five major areas of research.

	Number of researchers	Publications, full counts	Publication points, former	Publication points, new
Humanities	1,074	3,7	3,9	4,6
Social Sciences	1,882	4,5	3,1	4,7
Health Sciences	5,724	6,3	1,5	4,5
Natural Sciences	3,594	5,4	1,8	4,7
Engineering Sciences	2,157	5,5	2,1	4,4

The Norwegian model measures productivity as expressed in *publication points* at the level of *institutions*. Only peer-reviewed scholarly and scientific publications are included. In one dimension, three main publication types are given different weights: articles in journals and series, articles in books, and books. In another dimension, publication channels are divided into two levels in order to stimulate publishing in the most prestigious and demanding publication channels. We found that changing the weights in this system (see Schneider 2009, Sivertsen 2010, Ahlgren et al. 2012, or Aagaard et al. 2015 for detailed descriptions) *will not influence the balance*.

Instead, as expected from the evaluation, the problem of imbalance is directly related to the treatment of *co-authored* publications in the indicator. In the former indicator, the points for publications with multiple authors representing several institutions were *fractionalized* among the institutions according to their share of the authors. As seen in Table 1, this calculation creates an imbalance in favour of the humanities and social sciences. Applying full counts, on the other hand, creates an imbalance in the other direction. We found that the balance is reached by applying the *square root* of the institutional fraction of the publication, as seen in the column at the right.

It is important for this solution that the calculation is made from the point of view of the institution, not the author. There is no fractionalization of publications with authors from only one institution. On the other hand, an author with more than one affiliation will add to the denominator. The numerator of the institution's fraction is the number of authors from a particular institution. The denominator is the total number of unique combinations of authors and affiliations. The square root of an institution's fraction can never be more than one. On the other hand, the square root reduces the influence of fractionalization, thereby incentivising collaboration. We will discuss all issues related to the incentive structure of the new indicator in our final paper.

STUDY B: ADDING A MEASUREMENT OF CITATION IMPACT

The evaluation found that the citation impact of Norwegian research has remained stable in spite of a clear increase in publications. This effect is different from the decrease in citation impact that was experienced in Australia in a similar situation earlier on (Butler 2004). The design of the Norwegian indicator – with differentiated publication counts – seems to be the explanation for the different effects in the two countries (Schneider et al. 2016).

The present Norwegian government, however, has higher ambitions than maintaining stability. Study B recently started up to address the following question: Is it possible to supplement the publication indicator with a citation indicator and still provide a balanced representation of all fields? The answer will be given in our final paper, but some preliminary ideas can be presented here.

There are wide differences between fields of research with regard to the applicability and validity of citation indicators. The limitations mainly occur in the social sciences and humanities (Nederhof 2006), partly because the publication pattern itself has a limited coverage in citation databases such as Scopus and Web of Science (Sivertsen 2016b), partly because a limited number of references can be matched to source documents in the same databases (Sivertsen & van Leeuwen 2014). The latter limitation more specifically raises validity problems. Some fields have citation practices where only a smaller part of the references are given to recently published research. All such problems and limitations must be handled in a solution that should represent all fields in a balanced way.

The preliminary idea is to use a traditional citation indicator at the institutional level (field-normalized; weighted according to research profile; comparing to world average; based on all indexed publications or the ten per cent most cited) and apply it as a *multiplier* for the aggregated publication points at the institutional level. With an average citation rate at the aggregated level, the multiplier would be one. The *extent* to which the multiplier will be applied, will depend on three factors: the share of citable publications (coverage of the

institution's publications in Scopus or Web of Science); the number of citeable publications (a threshold is needed for reliability); and the number of publications in fields that are predefined as excluded because of validity problems. All three factors can be estimated annually (thereby dynamically) within the data for the publication indicator. In short, the idea is to let the citation indicator modify the publication indicator to the same extent, as it is applicable in the disciplinary continuum of different publication and citation practices.

REFERENCES

Aagaard K., Bloch C. W., Schneider J. W., Henriksen D., Lauridsen P. S., Ryan T. K. (2014). *Evaluation of the Norwegian Publication Indicator* (in Danish). Oslo: Norwegian Association of Higher Education Institutions.

Aagaard, K. (2015). How incentives trickle down: Local use of a national bibliometric indicator system. *Science and Public Policy*, 42(5), 725-737.

Aagaard, K., Bloch, C., & Schneider, J. W. (2015). Impacts of performance-based research funding systems: The case of the Norwegian Publication Indicator. *Research Evaluation*, 24(2), 106-117.

Ahlgren, P., Colliander, C., Persson, O. (2012). Field normalized citation rates, field normalized journal impact and Norwegian weights for allocation of university research funds. *Scientometrics*, 92(3): 767-780.

Bloch, C., & Schneider, J. W. (2016). Performance-based funding models and researcher behavior: An analysis of the influence of the Norwegian Publication Indicator at the individual level. *Research Evaluation*, doi:10.1093/reseval/rvv047.

Butler, L. (2004). What happens when funding is linked to publication counts? In H.F. Moed, W. Glänzel & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology* (pp. 340-389). Dordrecht: Kluwer Academic Publishers.

Nederhof, A. J. (2006). Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A Review. *Scientometrics*, 66(1), 81-100.

Piro, F. N., Aksnes, D. W., & Rørstad, K. (2013). A macro analysis of productivity differences across fields: Challenges in the measurement of scientific publishing. *Journal of the American Society for Information Science and Technology*, 64(2), 307-320.

Schneider, J.W. (2009). An outline of the bibliometric indicator used for performance-based funding of research institutions in Norway, *European Political Science* 8 (3): 364-378.

Schneider, J. W., Aagaard, K., & Bloch, C. W. (2015). What happens when national research funding is linked to differentiated publication counts? A comparison of the Australian and Norwegian publication-based funding models. *Research Evaluation*. DOI: 10.1093/reseval/rvv036

Sivertsen, G. (2010). A performance indicator based on complete data for the scientific publication output at research institutions. *ISSI Newsletter*, 6(1), 22-28.

Sivertsen, G., & Larsen, B. (2012). Comprehensive bibliographic coverage of the social sciences and humanities in a citation index: An empirical analysis of the potential. *Scientometrics*, 91(2), 567-575.

Sivertsen, G. & Schneider, J.W. (2012). *Evaluering av den bibliometriske forskningsindikator*. Oslo: NIFU.

Sivertsen, G. & van Leeuwen, T. (2014). Science, Technology & Innovation Indicators 2014. Thematic Paper 2. Scholarly publication patterns in the social sciences and humanities and their relationship with research assessment. Utrecht: Dialogic.

Sivertsen, G. (2016a). Publication-based funding: The Norwegian Model. In M. Ochsner et al. (Eds.), *Research Assessment in the Humanities*, in press, Springer.

Sivertsen, G. (2016b). Patterns of internationalization and criteria for research assessment in the social sciences and humanities. *Scientometrics*. DOI: 10.1007/s11192-016-1845-1

Technopolis Group. (2013). Evaluation of the Flemish Academic Bibliographic Database for the Social Sciences and Humanities (VABB-SHW). Executive Summary. Amsterdam: Technopolis Group.



Measuring research in Humanities and Social Sciences: information from a new Italian data infrastructure

Tindaro Cicero^{*}, Marco Malgarini^{**}

[*Tindaro.cicero@anvur.it](mailto:Tindaro.cicero@anvur.it)

Research Evaluation, ANVUR, Via Ippolito Nievo 35, 00153, Rome, Italy

[**marco.malgarini@anvur.it](mailto:marco.malgarini@anvur.it)

Research Evaluation, ANVUR, Via Ippolito Nievo 35, 00153, Rome, Italy

ABSTRACT

Measuring research output in Humanities and Social Sciences (HSS) is particularly important, since in these fields scientific production is much more heterogeneous than in Natural and Life Sciences, and as such it is not well represented in standard international databases normally used to assess research output and impact. For these reason, ANVUR has recently started a new data infrastructure, aimed at gathering information about scientific production, research infrastructures and research groups active in the Italian Universities. On the basis of these data, the aim of this paper is to provide a first characterization of Italian research Departments active in HSS, clustering them according to their level of research productivity and infrastructure availability. On the basis of our analysis, it is generally possible to distinguish among two main groups of Departments, respectively characterized by higher productivity but lower research quality, or by higher shares of excellent articles, but lower overall number of publications.

INTRODUCTION

The measurement of research quality and impact in Social Science and Humanities (HSS) is one of the most debated issues in STI studies (see among others Finkenstead, 1990; Hug et al, 2014; Ochsner et. al, 2012). The aim of our paper is to contribute to this discussion by exploiting a new data set on SSH research in Italy, based on the so-called Annual Synthetic Card for Departmental Research (Scheda Unica Annuale per la Ricerca Dipartimentale, henceforth SUA-RD). On the basis of these data, we aim at clustering Italian HSS Departments according to their scientific production, access to research infrastructures and participation in research groups.

THE DATA-SET

The main goal of SUA-RD is to help self-assessment, evaluation and accreditation of universities, in the framework of the Italian system of Quality Assurance (QA) and in full compliance with the European standard and guidelines (ENQA et al., 2015). Available data refers to the period 2011-2013; in the future the database will be updated yearly. SUA-RD is divided in three parts, respectively concerning objectives and management of Departments, their research results and Third Mission activities. In the first part, Departments describe their research policies and internal QA system, providing also information about active research groups. At this stage, Departments have also to enumerate research infrastructures, active researchers and administrative officials. The second part contains information regarding publications issued by Department members, including bibliographical references and

international co-authors (if any); the Department has also to report international mobility of members and the list of national and international competitive projects won, scientific premium achieved, existing fellowships in international scientific societies, directions of scientific journals and book series, international research or teaching appointments and responsibilities for the organization of international conferences or workshops. Finally, the third part is devoted to monitoring Third Mission activities, including patents, spin-offs, third-party and public engagement activities. In the following we will concentrate upon information concerning the number and characteristics of publications, the number of research groups, available infrastructures and the composition of research staff.

RESEARCH STAFF AND ACADEMIC PRODUCTION IN HSS

Table 1. Number of research staff in the Italian Higher Education System in HSS.

Area	Full Professors	Associate Professors	Researchers	Phd Student	Other	Total
Architecture	422	574	1,000	1,638	501	4,135
Arts and Humanities	1,210	1,485	2,357	2,822	696	8,570
History and Philosophy	915	946	1,430	1,830	525	5,646
Law	1,488	1,107	2,130	2,816	570	8,111
Statistics and Economics	1,417	1,305	2,018	1,935	701	7,376
Political and Social sciences	385	446	876	999	387	3,093
Total	5,837	5,863	9,811	12,040	3,380	36,931

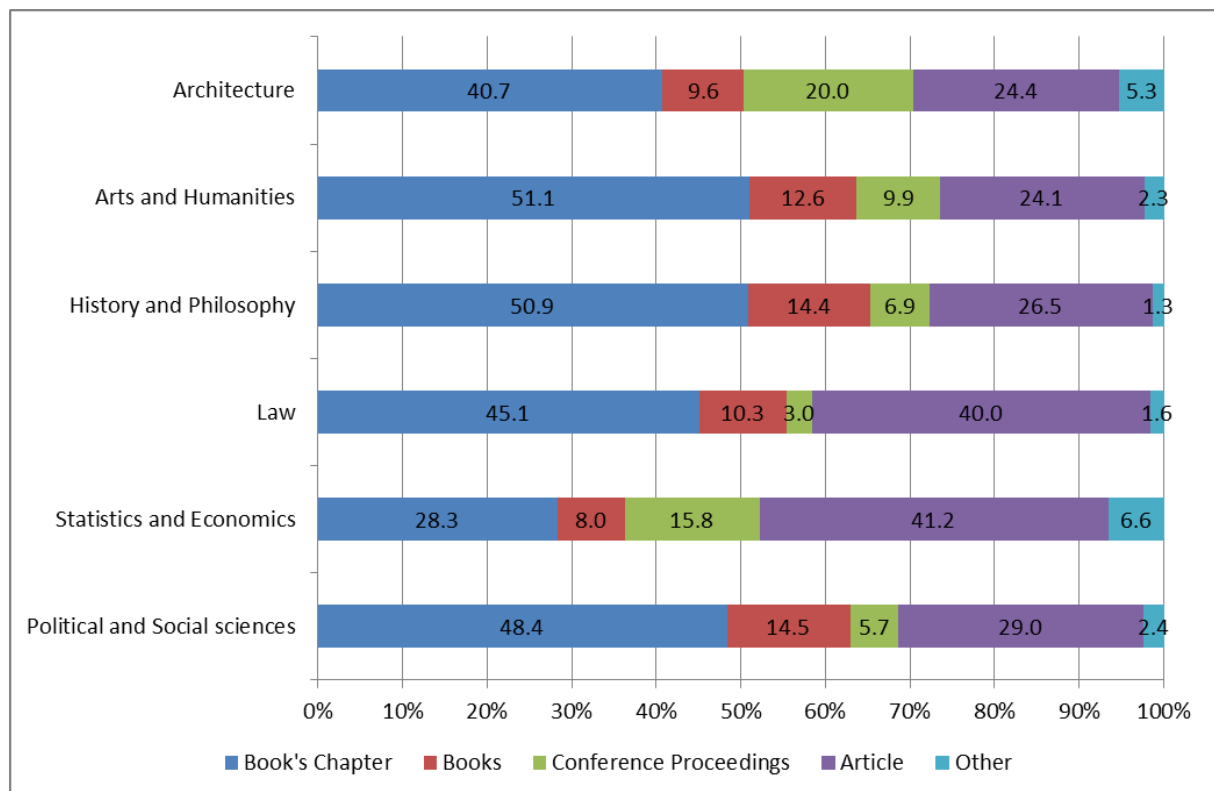
Data about scientific production is referred to full and associate professors, assistants and PhD students, with the exclusion of students specializing in medical areas (Table 1). Only scientific publications are considered, i.e. excluding those merely educational or informative, according to the authors themselves. Total scientific production in Italian HSS for the years 2011-2013 has been equal to roughly 235,000 items, including articles, conference proceedings, books, books chapters and other publications (i.e. patents, design, posters, etc.; see Table 2). Arts and Humanities show the largest scientific production with 57,514 research products, while Political and Social sciences report only 19,349 publications.

Table 2. Academic production in HSS by publication year and area.

Area	2011	2012	2013	Total
Architecture	8,661	10,805	9,148	28,614
Arts and Humanities	19,165	20,631	17,718	57,514
History and Philosophy	13,467	14,315	13,893	41,675
Law	16,429	18,026	16,148	50,603
Statistics and Economics	11,640	13,360	11,620	36,620
Political and Social sciences	6,288	6,777	6,284	19,349
Total	75,650	83,914	74,811	234,375

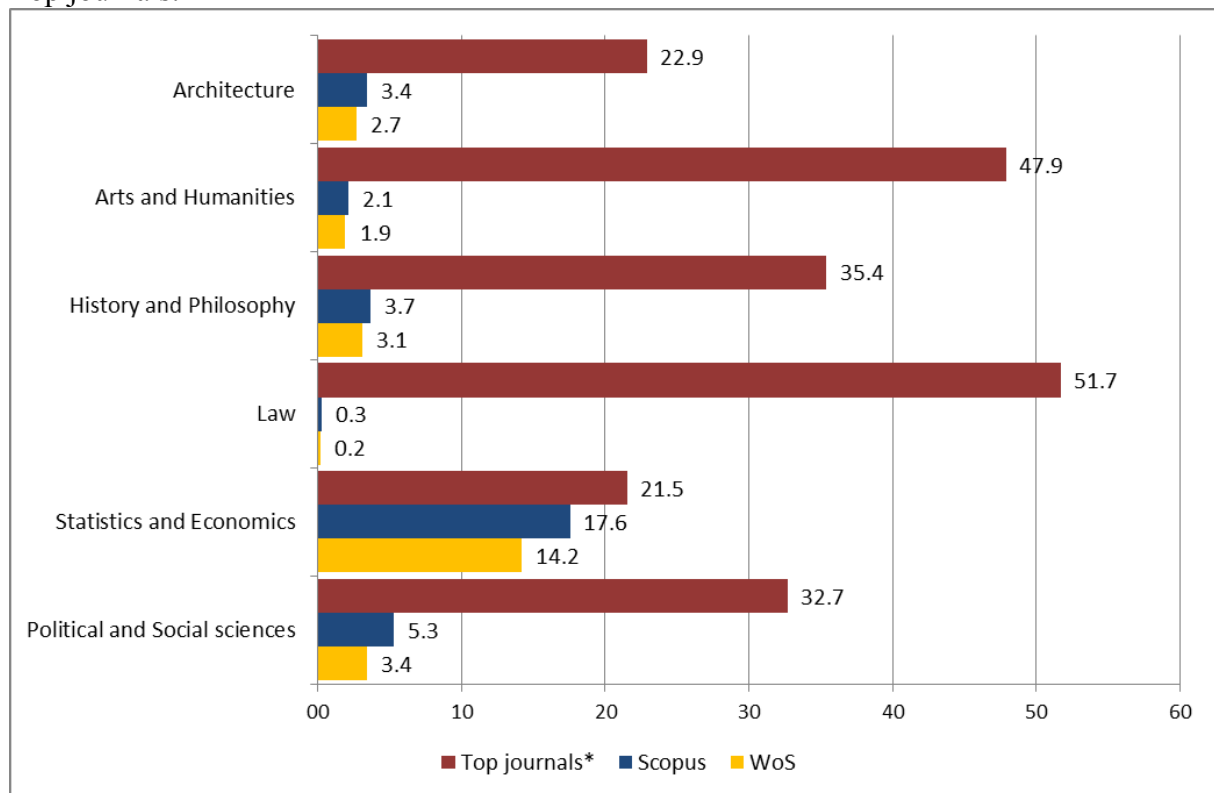
Books chapters are the commonest way to share knowledge in HSS; they are particularly important in Arts & Humanities and in History and Philosophy. Articles are especially relevant in Statistics and Economics and in Law, while Conference Proceedings are more significant in Architecture.

Figure 1. Academic production in HSS by typology and area.



The share of articles covered in the two most important international databases (Web of Science and Scopus) ranges from a maximum of roughly 15% (depending on the database considered) in Economics and Statistics to a minimum of less than 1% in Law (see Figure 2). This evidence supports the view that available databases are not sufficiently representative to evaluate research in HSS (Van Leeuwen, 2013); for this reason, since 2012 ANVUR has asked a group of experts to identify among the major national and international journals those to be considered as scientific (as distinguished from those purely educational or informative) and, among them, those deemed as excellent in term of international impact and esteem and for the rigor of their referee process (Bonaccorsi et al., 2015). The percentage of Top, excellent journals varies across areas, with a peak in Law (51.7%).

Figure 2. Articles in HSS: coverage in international databases and percentage published in Top journals.



RESEARCH FACILITIES, LIBRARIES AND RESEARCH GROUPS IN HSS

Large research infrastructures represent an important tool for developing Universities strategy and supporting research activities. With the aim of encouraging the dissemination of good research infrastructures, the European Commission in 2002 set the European Strategy Forum for Research Infrastructures (ESFRI), aimed at defining the needs for infrastructure and at implementing a strategic roadmap for their construction. For Italy, the SUA-RD provides a first census of all major research infrastructures used in Italian Universities. In SUA-RD, large Research Facilities are defined as those characterized by a value at constant prices higher than 100.000 Euro and a high degree of specialization. Among those, we can find a heterogeneous range of facilities pertaining to different research fields, from medical sciences to Human and Social Sciences. Large Research Facilities (LRFs) operating in HSS represent 5% of the total number of LRFs censused in the SUA-RD; a majority of them are related to archeological activities. Thirty-five Universities have at least one large facility, with some differences emerging at the local level (Table 3).

Table 3. Large Research Facilities classified in the Human and Social Sciences by geographic area.

Geographic area	N. of Large Research Facilities	N. of supervisors	N. of supervisors for LRI
North-East	30	43	1.4
North-West	20	23	1.1
Centre	59	82	1.4
South	84	110	1.3
Islands (Sicily and Sardinia)	17	28	1.6
On-line universities	1	1	1
Italy HSS	211	287	1.4

Research groups are supposed to enhance scientific quality (Andrade et al., 2009); unfortunately, data about them is usually scarce, and as a consequence some studies tried to identify them using co-authorship networks (Perianes et al., 2010). In this sense, the SUA-RD provide a unique source of information in order to study the relationship among scientific productivity and participation to such groups; in SUA Rd, a research group is identified as a set of more than one researchers working together (formally or informally) on a specific project or program of research. For each research group the SUA-RD gathers information about the objectives, lines of research and staff composition (internal and external). In HSS, research groups are usually associated with knowledge transfer activity (Olmos-Peñuela et al., 2014). Research groups censused in the Italian HSS sectors are 3,401 (see Table 4). The average number of component for group is 5.3 and there are not significant differences among geographic areas.

Table 4. Number of research groups in HSS

Geographic Area	N° Research Groups	N° of internal components	N° of external components	Average number of components for group
North-East	829	3,782	534	5.2
North-West	631	3,403	299	5.9
Centre	747	3,471	346	5.1
South	819	3,809	406	5.1
Islands (Sicily and Sardinia)	354	1,675	165	5.2
On-line universities	21	60	15	3.6
Italy HSS	3,401	16,200	1,765	5.3

Table 5 reports the number of libraries, books and journals available for research in HSS; recent literature (Haglund and Olsson, 2008) has pointed out that IT progress has reduced the need of consulting physical libraries; nonetheless, availability of bibliographic resources is supposed to remain an important factor driving the production of high quality research.

Table 5. Libraries in HSS

Geographic Area	N° of libraries	N° of books	N° of journals
North-East	105	13,138,644	4,244,605
North-West	144	10,788,102	2,773,840
Centre	208	24,442,928	7,740,159
South	185	6,818,247	1,740,791
Islands (Sicily and Sardinia)	89	4,337,856	1,333,412
On-line universities	13	12,313	935
Italy HSS	744	59,538,090	17,833,742

CLUSTERING ITALIAN UNIVERSITY DEPARTMENTS

SUA RD data may be used to provide a first classification of University Departments operating in HSS on the basis of their scientific output, their access to infrastructures and libraries and the participation to research groups. In the exercise, we consider a total of 803 Department/Area couples with at least 5 persons per Department/Area. We use cluster analysis in order to classify these Departments in homogenous groups; groups are chosen so as to minimize within-group and maximize between-groups variance. For each Area, the optimal number of clusters is determined according to the Calinsky/Harabasz (1974) pseudo F stopping rule¹.

In Law, History and philosophy and Political and social sciences we are able to distinguish five clusters of Departments; in Economic and statistics, we identify only two clusters, while in Architecture we identify three clusters and in Arts and humanities, four (Table 5). More specifically, in Law and Political and social sciences a first group of Departments show a high number of publications, most of them of remarkable scientific profile. In Political and social sciences this group is composed by 12 departments and it is also characterized by an intensive production of articles and by the availability of libraries and infrastructures. In Law, the group is composed by 30 departments and is characterized by relevant participation to research groups and high number of Professors. In History and philosophy, the group of high quality departments is also the most productive and is especially concentrated in the production of books chapters. In Economics and statistics, we identify a first group of Departments characterized by a high number of publications per capita but a low share of scientific and outstanding products, and another with high-quality research and a reduced overall number of publications. Access to infrastructures and participation to research groups are associated with a higher number of publications. In Architecture we distinguish three clusters: a first group of Departments is characterized by high number of conference proceedings and frequent access to libraries; a second group is characterized by high numbers of publications per capita but a low share of outstanding publications, whilst a third group is defined by its high-quality research profile, associated with a high share of full professors.

¹ Clusters were selected with the K-medians cluster method and the Canberra (dis)similarity measure, using the software Stata 13 ©.

Table 6. Italian Departments' Clusters

Clusters	N	1	2	3	4	5	6	7	8	9	10	11	12
Architecture													
1	18	0.290	0.207	0.357	0.094	0.057	0.151	0.083	0.154	0.073	6.007	0.076	7.096
2	23	0.227	0.233	0.394	0.098	0.054	0.184	0.095	0.225	0.000	1.176	18.633	9.044
3	22	0.254	0.236	0.366	0.087	0.069	0.192	0.130	0.144	0.847	3.994	9.362	7.649
Total	63	0.255	0.227	0.374	0.093	0.060	0.177	0.104	0.177	0.317	3.540	10.094	8.000
Arts and Humanities													
1	33	0.101	0.277	0.458	0.146	0.129	0.225	0.141	0.186	0.445	0.057	0.034	6.960
2	20	0.145	0.214	0.506	0.121	0.089	0.184	0.154	0.200	2.676	11.138	2.547	7.362
3	41	0.106	0.248	0.492	0.126	0.120	0.217	0.153	0.180	0.404	4.644	15.703	7.539
4	57	0.118	0.226	0.499	0.135	0.105	0.195	0.151	0.194	0.018	0.000	18.149	7.215
Total	151	0.115	0.241	0.489	0.133	0.112	0.206	0.150	0.189	0.568	2.749	11.460	7.267
History and Philosophy													
1	19	0.054	0.314	0.477	0.146	0.140	0.260	0.110	0.145	0.058	0.075	6.686	6.165
2	21	0.077	0.222	0.531	0.151	0.076	0.173	0.173	0.202	1.531	9.132	2.251	8.860
3	38	0.066	0.261	0.503	0.159	0.092	0.210	0.159	0.199	0.000	0.000	0.000	8.038
4	42	0.086	0.229	0.504	0.169	0.081	0.179	0.204	0.195	1.349	0.566	24.631	8.507
5	46	0.095	0.280	0.490	0.120	0.099	0.224	0.144	0.175	0.237	7.173	15.849	7.891
Total	166	0.079	0.259	0.500	0.148	0.094	0.207	0.163	0.185	0.607	3.295	11.674	8.005
Law													
1	54	0.039	0.391	0.448	0.109	0.195	0.353	0.183	0.180	0.000	2.110	0.000	6.860
2	15	0.000	0.420	0.437	0.133	0.226	0.403	0.229	0.196	0.000	2.730	1.684	5.201
3	30	0.024	0.438	0.414	0.101	0.247	0.413	0.229	0.203	2.292	0.392	23.056	7.842
4	18	0.053	0.311	0.491	0.133	0.141	0.282	0.177	0.153	0.047	6.688	11.663	6.868
5	37	0.036	0.404	0.439	0.113	0.194	0.358	0.172	0.163	0.386	3.820	12.562	6.182
Total	154	0.033	0.397	0.443	0.113	0.201	0.362	0.193	0.179	0.545	2.782	9.037	6.728
Economics and Statistics													
1	79	0.157	0.419	0.268	0.088	0.146	0.378	0.206	0.188	0.199	3.280	0.000	6.585
2	82	0.178	0.402	0.283	0.081	0.133	0.366	0.197	0.187	0.235	5.217	19.964	6.777
Total	161	0.168	0.410	0.276	0.084	0.139	0.372	0.201	0.187	0.218	4.266	10.168	6.683

	6												
Political and Social Sciences													
1	40	0.055	0.290	0.497	0.143	0.084	0.215	0.131	0.175	0.000	2.313	0.000	7.109
2	14	0.071	0.299	0.457	0.163	0.098	0.220	0.085	0.078	0.000	1.263	8.754	4.640
3	14	0.082	0.218	0.516	0.158	0.038	0.163	0.102	0.174	0.056	8.935	13.689	7.674
4	28	0.040	0.286	0.482	0.153	0.092	0.228	0.170	0.158	0.510	0.227	22.066	6.647
5	12	0.031	0.331	0.460	0.161	0.132	0.254	0.182	0.159	1.667	10.047	16.424	8.549
Total	108	0.054	0.286	0.486	0.152	0.087	0.217	0.137	0.156	0.325	3.354	10.455	6.903

(N) Number of observations; (1) Percentage of proceedings; (2) Percentage of articles; (3) Percentage of books chapters; (4) Percentage of books; (6) Percentage of scientific articles; (5) Percentage of excellent articles; (7) Percentage of Full Professors; (8) Percentage of Associate Professors; (9) Infrastructures per 100 persons; (10) Libraries per person; (11) Research groups per person; (12) Publications per person.

Finally, in Arts and humanities we identify four groups of similar size: 1) Departments with high quality research, but low productivity and a few number of available libraries and infrastructures; 2) Departments characterized by frequent access to infrastructures, high number of full and associate professors and a scientific activity particularly concentrated in the production of conference proceedings and books chapters; 3) Departments with high productivity of outputs with a medium-low quality profile; 4) Departments characterized by a strong presence of research groups.

CONCLUDING REMARKS

The data provided in the SUA-RD represents a new, primary source of information on research carried by Italian Universities. In this paper, we have concentrated our attention on Departments active in HSS, clustering them on the basis of their research activities, access to research physical and bibliographical infrastructures and participation to research groups. Generally speaking, we manage to distinguish at least among two main groups of Departments, one characterized by high productivity but research quality lower than the average, and the other by a higher share of excellent articles, but a lower number of publications. In some field, however, we are also able to further distinguish other groups characterized by both low levels and low quality of production. Political and Social Sciences and Law Departments stand out, in the sense that the most productive departments in these fields are also those characterized by higher quality of production.

The analysis presented in the paper represents a first exploration of this new, large dataset. Further research is advisable, trying in particular to better understand the possible determinants of different research results obtained by the Departments, trying for instance to incorporate data on University financing or looking more deeply into the composition of the research staff (considering for instance gender effects and the impact of academic experience) and at possible effects of network externalities linked to the geographical location and size of the University.

REFERENCES

- Andrade, H. B., de Los Reyes Lopez, E., & Martín, T. B. (2009). Dimensions of scientific collaboration and its contribution to the academic research groups' scientific quality. *Research Evaluation*, 18(4), 301-311.
- Bonaccorsi, A., Cicero, T., Ferrara, A., & Malgarini, M. (2015). Journal ratings as predictors of articles quality in Arts, Humanities and Social Sciences: an analysis based on the Italian Research Evaluation Exercise. *F1000Research*, 4.
- Calinski, T., and J. Harabasz. (1974), "A dendrite method for cluster analysis." *Communications in Statistics*. Vol. 3, No. 1, pp. 1–27.
- European Association for Quality Assurance in Higher Education (ENQA), European Students' Union (ESU), European University Association (EUA), European Association of Institutions in Higher Education (EURASHE), (2015), Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG), Bruxelles, Belgium, http://www.enqa.eu/wp-content/uploads/2015/11/ESG_2015.pdf
- Finkenstaedt, T. (1990). Measuring research performance in the humanities. *Scientometrics*, 19(5/6), 409–417.
- Haglund, L., & Olsson, P. (2008). The impact on university libraries of changes in information behavior among academic researchers: a multiple case study. *The journal of academic librarianship*, 34(1), 52-59.
- Hug, S. E., Ochsner, M. & Daniel, H.-D. (2014). A framework to Explore and Develop Criteria for Assessing Research Quality in the Humanities. *International Journal for Education Law and Policy*. Forthcoming
- Olmos-Peñuela, J., Castro-Martínez, E., & D'Este, P. (2014). Knowledge transfer activities in social sciences and humanities: Explaining the interactions of research groups with non-academic agents. *Research Policy*, 43(4), 696-706.
- Ossenblok, T. L., Engels, T. C., & Sivertsen, G. (2012). The representation of the social sciences and humanities in the Web of Science—a comparison of publication patterns and incentive structures in Flanders and Norway (2005–9). *Research Evaluation*, 21(4), 280-290.
- Perianes-Rodríguez, A., Olmeda-Gómez, C., & Moya-Anegón, F. (2010). Detecting, identifying and visualizing research groups in co-authorship networks. *Scientometrics*, 82(2), 307-319.
- Van Leeuwen, T. (2013). Bibliometric research evaluations, Web of Science and the Social Sciences and Humanities: a problematic relationship?. *Bibliometrie-Praxis und Forschung*, 2.



Trends and Developments in Multi-Authorship in five Social Science disciplines from 1991 to 2014

Sabrina J. Mayer *

* mayer@forschungsinfo.de

German Centre for Higher Education Research and Science Studies (DZHW), Department 2, Schützenstraße 6a, 10119 Berlin (Germany)

ABSTRACT

This article explores developments in author numbers in five disciplines of the Social Sciences from 1991 to 2014, Economics, Educational Sciences, Political Science, Psychology, and Sociology. The relationship between the mean number of authors per article (in every discipline's five top journals) and three factors, publication year, international collaboration (IC), and article content (theoretical, experimental, or large-scale comparative) are analyzed.

The findings show that single-authorship is decreasing and multi-authorship is on the rise, publication year was found to be significantly correlated with the number of authors. In addition, IC is positively correlated with the number of authors, even when publication year is controlled for. The content type “theoretical articles” is negatively related to the number of authors, even when publication year is controlled for.

Differences between the disciplines were found for the development of multi-authorship: Psychology and Educational Sciences tend to have high shares of co-authored articles whereas single authorship is still dominant in Political Science and Sociology. In addition, differences also exist for the relationship of article content and the number of authors: Experimental articles have on average more authors than large-scale/comparative articles in Economics, Political Science and Sociology, whereas large-scale/comparative articles have on average more authors in Psychology.

This article's significant findings show that aggregating disciplines from the Social Sciences may not always be appropriate, as differences in authorship patterns exist that could not be observed without separate analyses. In addition, articles' content types are found to be correlated with author numbers, as it was already known from earlier studies of the sciences, so this dimension should be considered in further analyses.

INTRODUCTION

One of the most extreme examples for multi-authorship was set in May 2015 when an article in *Physical Review Letters* listed 5,154 authors (Aad et al. 2015). But increasing author numbers is not a trend singular to Physics: A steady rise of the mean number of authors per article can be observed in almost all disciplines (Cunningham & Dillon, 1997; Glänzel, 2002; Hudson, 1996; Levsky, Rosin, Coon, Enslow, & Miller, 2007; Sin, 2011). Two motives can be named for this increase in authorship numbers. First, the way of knowledge production has changed. The gradual change of the organization of science, named by Price (1986: 2-4) as the transition from Little to Big Science, affects the number of authors per paper. Whereas short-time projects, small teams, without specialized members were characteristic for Little Science, Big Science includes large-scale, often external funded projects, with specialized members in huge, often international, teams (Price 1986: 7, 59-61, 77; Chompalov, 2014). This “general movement towards mass collaboration” (Price, 1986: 77), that already started, albeit at a low level, at the beginning of the 20th century, is related to the exponential growth that is observed for all parts of science (1986: 79). Recent studies on the organization of science (Gibbons et al., 1994; Nowotny, Scott, & Gibbons, 2001) built on this distinction and identified two different paradigms of scientific production: “Mode 1” as a traditional, disciplinary-based way of knowledge production, which is slowly replaced by a new paradigm, “Mode 2”, where knowledge production is global, project-oriented, and transcends old disciplinary borders. In contrast to Price who mainly referred to homogenous output-related growth within a field, they focus on heterogeneous growth as a process of differentiation (Gibbons, Limoges, & Nowotny et al., 1994). This does not only increase the number of authors but also the number of (international) collaborations: “Not only is the average number of authors per paper increasing [...] but the geographical distribution of these institutions continues to broaden.” (Gibbons, Limoges, & Nowotny et al., 1994: 34)

Second, developments in politics and society are another driving force for the changing authorship pattern in academic publication. The shift towards entrepreneurial universities and the rise of managerial governance lead to greater importance of output-related indicators that have become crucial for obtaining funding (Hattke, Blaschke, & Frost, 2014; Slaughter & Leslie, 1997). In addition, funding for research can be obtained easier by international teams nowadays (e.g. most Horizon 2020 funding opportunities support proposals by institutions from at least three different countries).

Analyses of authorship patterns were already conducted for different academic fields, disciplines or specific countries but focus mostly on STEM disciplines and the life sciences (Cunningham & Dillon, 1997; Glänzel, 2002; Glänzel & Schubert, 2004; Hudson, 1996; Levsky, Rosin, & Coon et al., 2007; Sin, 2011). Studies of authorship patterns in the Social Sciences either analyze only a single subject (e.g. Hudson, 1996; Kliegl & Bates, 2011), or provide only aggregated analyses (Endersby, 1996; Engels, Ossenblok, & Spruyt, 2012; Leydesdorff, Park, & Wagner, 2014; Nederhof, 2006). Both approaches are problematic: Single analyses do not allow a comparative perspective and aggregated analyses blur the differences between the disciplines. In addition, the existing studies often describe authorship patterns but do not try to explain these developments (in a multi-variate way) (e.g. Endersby, 1996; Huang, 2015; Hudson, 1996).

RESEARCH QUESTION AND HYPOTHESES

This article aims to provide a comparative, longitudinal perspective on current trends in authorship patterns in the Social Sciences from 1990-2014 for five disciplines, Economics, Educational Sciences, Political Science, Psychology, and Sociology. Three factors were identified that could explain developments in authorship numbers. As a time period of 24 years is analysed in this study, a general driving force could actually be the publication year

of an article. Price (1986) named science's exponential growth as a reason for his now famous prediction that "by 1980 the single-author paper will be extinct" and scientific publications may „move steadily toward an infinity of authors per paper" (1986: 79). Therefore, there should be a positive relationship between the mean number of authors per article and the year of publication (H1a) In addition, the share of co-authored papers should be higher in all disciplines now than at the beginning of the analysed time period (H1b).

Two further factors are analysed to explain rising authorship numbers, international collaborations (IC) and article content. First, international collaborations are more likely to be research with larger levels of division of tasks, therefore IC and the average number of authors per paper should have a positive relationship (H2a). However, this factor is surely be confounded with time: In the last years, IC were on the rise because of the changing ways of knowledge production, more cooperation and multi-national teams are a consequence of Mode 2 as well as of the transformation to Big Science, and were facilitated by sinking communication costs (Glänzel, 2002; Hoekman, Frenken, & Tijssen, 2010; Sin, 2011). Therefore, multi-variate analyses are needed to show that the relationship of IC and the number of authors is not caused by the time-specific general trend. It is supposed that there is a significant correlation between IC and the number of authors, even when the effect of the year of publication is controlled for (H2b).

Second, the relationship between author numbers and article content is analysed. It is already known from the sciences that theoretical articles have fewer authors on average than experimental articles (e.g. Katz & Martin, 1997). Analogously, it is expected that theoretical articles have fewer authors than other articles as it is often (H3a). Experimental work in the Social Sciences is different from other sciences and has a varying standing in the different disciplines. In some disciplines, lab experiments are a common form of research as in Educational Sciences and Psychology (e.g. Gorard, 2004; Martin & Sell, 1979). In other disciplines such as Political Science and Sociology, experiments became popular only recently and often involve large field experiments that need more resources than traditional lab experiments (Eifler, 2014; Gerber & Green, 2013). Concerning studies of voting behaviour or value orientations in different countries, data for large-scale comparative studies are easily available online (e.g. the World Value Survey, the Comparative Study of Electoral Systems or statistical data on households and companies), whereas cross-national data are scarce in Educational Sciences and Psychology. Collecting comparative data is often an expensive endeavour in these disciplines that needs international teams and external funding. Therefore, it is expected that large-scale/cross-national studies in Economics, Political Science and Sociology have fewer authors than experimental works (H3b). Contrary, experimental works in Educational Sciences and Psychology are assumed to have fewer authors (H3c). However, this factor may also be confounded with time. It is supposed that the differences between author numbers based on content remain significant, even when time is controlled for (H3d).

DATA SOURCE AND METHODOLOGY

Data is drawn from an in-house Web of Science (WoS) data base by Thomson Reuters. Only publications with the doctype "article" are included in this study. As this paper aims to provide a comparative analysis by discipline, only discipline-specific journals were chosen. For a data set that is insensitive to outliers, five general journals with the highest impact factor for each discipline, going back at least until 1980, were pooled for the analyses (see Table 1).

Table 1: Analysed journals, and the number of articles used for the study 1991-2014

Journal Title	Earliest article available in WoS	Number of analysed articles
Economics		
Journal of Economic Literature	1969	132
Quarterly Journal of Economics	1902	946
Journal of Finance	1951	1,710
Journal of Financial Economics	1979	1,729
Review of Economic Studies	1936	973
Educational Studies		
Review of Educational Research	1933	201
Medical Education	1978	2,489
American Educational Research Journal	1966	676
Journal of Learning Disabilities	1970	1,047
Journal of Teacher Education	1975	451
Political Science		
American Political Science Review	1906	1,011
International Organization	1969	520
American Journal of Political Science	1973	1,292
Policy Studies Journal	1975	805
Journal of Peace Research	1967	828
Psychology		
Psychological Bulletin	1904	426
Neuroscience And Biobehavioral Reviews	1979	430
Psychological Review	1900	538
Journal of Personality And Social Psychology	1965	3,457
Journal of Child Psychology And Psychiatry	1966	1,880
Sociology		
American Sociological Review	1936	1,079
American Journal of Sociology	1900	789
Sociological Methods & Research	1973	451
Sociological Quarterly	1966	779
Sociology-The Journal of The British Sociological Association	1970	1,094

The use of bibliometric indicators in the Social Sciences can be problematic because research is still more often published in books/book chapters than in other disciplines (Hicks, 1999; van Leeuwen, van Wijk, & Wouters, 2016).¹ As journal articles are on the rise in the Social Sciences, the coverage gap decreases (Engels, Ossenblok, & Spruyt, 2012; Leydesdorff, Park, & Wagner, 2014). However, there may be, at least in some disciplines, a divide in the academic community between authors that still publish monographs and others that mainly rely on journal articles (Cronin, Snyder, & Atkins, 1997).² This could lead to the fact that the

¹ Based on an analysis of the output of one Dutch university 2004-2009, van Leeuwen, van Wijk, & Wouters, (2016) report that 36.8 percent of all publications in the Social Sciences (without Economics) are journal publications, and the internal coverage of Social Sciences' journal articles in WoS is about 61 percent.

² I would argue, contrary to Cronin et al. (1997), that both groups may probably have converged over time, and that new discoveries are now first published as journal articles, whereas book chapters/monograph often follow and deals with already partly published results. So far, there are no recent studies on this.

data only allows the analysis for a particular group of researchers that belong to the second group.

Disciplines were coded according to the OECD knowledge fields using the matching of subject categories provided by Thomson Reuters. However, in the OECD coding, Economics and Business are grouped together although both are highly different (e.g. Hudson, 1996), instead, the WoS subject category “Economics” was used. International collaboration is operationalized as articles from two or more authors coming from more than one country. All articles without a valid address (institution and country) were excluded (n=355).

For the analysis of an article’s content, a textual analysis of the articles’ abstracts is performed with regular expressions³ in R. Articles are classified into four content types, theoretical, experimental, large-scale/comparative or other articles. Articles are only classified as theoretical work if theoretical terms are matched and no matches are found for empirical terms to ensure that theory-based empirical studies are excluded. The resulting numbers of classified articles are 1,808 (theoretical), 2,063 (experimental) and 3,284 (large-scale/cross-national work). 170 articles that were classified both as experimental and large-scale articles were excluded. No abstracts are available for articles before 1991, so the dataset encompasses 24,786 articles (see Table 2).

Table 2: Number of articles with abstracts by time period and discipline

	1991-94	1995-99	2000-04	2005-09	2010-14
Economics	566	975	1135	1316	1498
Educational Sciences	613	873	1155	1145	1078
Political Science	522	907	904	1030	1093
Psychology	745	1361	1269	1357	1509
Sociology	501	764	707	802	929

The maximum number of authors are 8 (Economics), 29 (Educational Sciences), 13 (Political Science), 50 (Psychology), and 12 (Sociology).

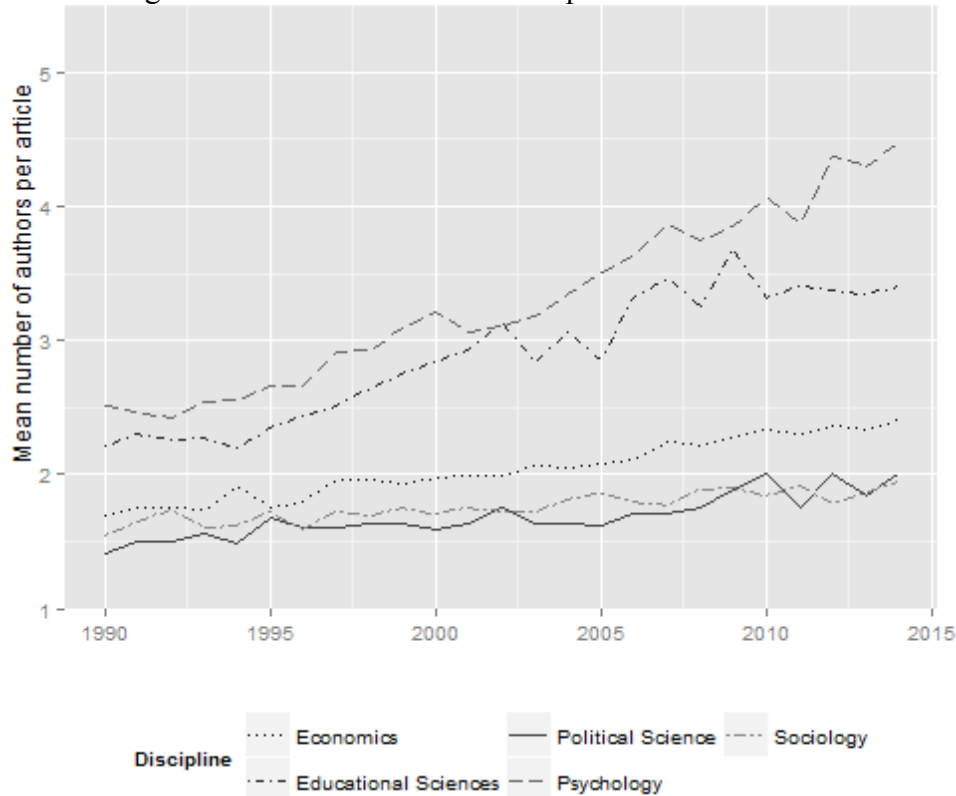
RESULTS

Relation between publication year and author numbers

First, the development of the mean number of authors per article by discipline over time is displayed in Figure 1. In all five disciplines, author numbers per paper increased between 1991 and 2014, however, the increase was higher in Psychology and Educational Sciences as it was in Political Science and Sociology. For every discipline, there is a low but positive correlation between publication year and the number of authors, ranging from $r= 0.08^{***}$ (Sociology) up to $r= 0.26^{***}$ (Psychology). H1a can be accepted.

³ Used expressions: Theoretical works: “theory|theoretical|argument |framework”
 Experimental works: “experiment |quasi-experiment|randomized”
 Large-scale: “longitudinal|nations|countries|different languages”

Figure 1: Mean number of authors per article 1991-2014



Next, the general pattern of co-authorship is explored (see Table 3). The development of shares of co-authored articles differs by discipline as well: More than 80 percent of all analysed articles in Economics, Educational Sciences and Psychology are now co-authored articles. Contrary, almost half of all articles are still single-authored in Political Science and Sociology.

Table 3: Mean number of authors and share of co-authored articles by discipline

	1991-94	1995-99	2000-04	2005-09	2010-14
Economics					
mean	1.77	1.89	2.02	2.19	2.35
% co-authored	60%	65%	72%	78%	81%
Educational Sciences					
mean	2.25	2.54	2.96	3.31	3.37
% co-authored	62%	70%	78%	82%	83%
Political Sciences					
mean	1.49	1.63	1.65	1.73	1.92
% co-authored	37%	45%	47%	52%	59%
Psychology					
mean	2.50	2.86	3.18	3.72	4.21
% co-authored	79%	85%	88%	93%	95%
Sociology					
mean	1.63	1.70	1.75	1.84	1.88
% co-authored	47%	49%	50%	53%	56%

However, the share of co-authored articles increased from 1991-1994 to 2010-2014 in all disciplines, ranging from 8.50 (Sociology) to 22.0 percent (Political Science). All mean

differences between these two time periods are significant at the 0.1 percent level.⁴ Therefore, H1b can be accepted.

Analysing co-authorship numbers by collaboration type

Next, the relationship between author numbers and cooperation type is analysed. It was expected that IC and the average number of authors are positively correlated. For all five analysed disciplines, moderate positive correlations were found, ranging from $r=0.22^{***}$ (Educational Sciences) to $r=0.32^{***}$ (Economics). The mean number of authors for IC is on average 1.21 ($p<0.001$) higher than for non-IC. One could argue that articles with one author can never be written in IC. Therefore, including single-authored articles in this analysis could lead to an over-estimation of the correlation between IC and the number of authors. If the analyses are repeated for all articles with more than one author, still highly significant, albeit lower positive correlations are found ($r=0.15^{***}$, Political Science to $r=0.21^{***}$, Psychology). The mean number of authors for IC is on average still 0.64 ($p<0.001$) higher than for other articles. H2b can be confirmed.

However, the share of articles written in IC is significantly increasing for all five disciplines from 1990-1994 to 2010-2014 (8, Sociology to 24 percentage points, Economics, all differences $p<0.001$). To allow an independent examination of the relationship of IC and the number of authors, partial correlation analysis is used to see if the relationship is caused by publication year⁵. However, even when publication year is controlled for, moderate, positive correlations between IC and the number of authors for all articles with more than one authors can be observed ($r=0.12^{***}$, Sociology to $r=0.18^{***}$, Psychology). H2b can be accepted.

Analysing co-authorship numbers by article content

Last, the relationship between article content and author numbers is analysed (see Table 4).

Table 4: Average number of authors per article content type and discipline

	Article content			
	All other	Theoretical	Experimental	Large-scale /comparative
Economics	2.09	2.05	2.55	2.20
Educational Sciences	3.08	2.33	3.35	3.33
Political Science	1.71	1.57	2.15	1.72
Psychology	3.45	2.69	3.05	4.00
Sociology	1.77	1.55	2.33	1.89
Total	2.53	1.96	2.88	2.55

Theoretical articles have the lowest average number of authors, ranging from 1.6 in Political Science and Sociology up to 2.7 in Psychology. The mean differences in author numbers between theoretical articles and the other three categories (all others, experimental and large-scale) are all significant at 0.1% level, except for Economics (theoretical vs. all other articles, $p<0.852$) and Psychology (theoretical vs experimental articles, $p<0.073$). H3a can be accepted. For experimental and large-scale/cross-national articles, consequences for the

⁴ For mean differences within disciplines, either t-tests or one-way anova with Scheffé post-hoc tests to control for alpha inflation, were performed. Because of the limited space, only p-values are reported.

⁵ As partial correlation analysis has the implicit idea of fitting a (linear regression) model to the data, with the number of authors as the dependent variable, one could argue that using Poisson/negative binomial regression analysis would be more appropriate. However, this leads to the same results, for each discipline: IC has its own effect on the number of authors, even when negative binomial regression analysis is used and publication year is controlled for.

number of authors are supposed to vary by discipline. Indeed, experimental articles have on average more authors in Economics (0.35), Political Science (0.33) and Sociology (0.44) than large-scale empirical articles (differences between both types of article content, $p < 0.001$). H3b can be confirmed. Contrary, large-scale empirical works in Psychology have on average almost one author more than experimental works ($p < 0.001$). No difference in average author numbers could be found for Educational Sciences. This is probably due to the highly different branches of the discipline (e.g. medical education and critical pedagogy). H3c can only be confirmed for Psychology. This factor may also be confounded with the publication year, so negative binomial regression⁶ is used to control for the effect of time simultaneously (see Table 5).

Table 5: Negative binomial regression on the number of authors, Regression coefficients

	Number of authors				
	Economics	Educational Sciences	Political Science	Psychology	Sociology
Content (Ref. Group: Experimental articles)					
All other articles	-0.160*** (0.05)	-0.070 (0.04)	-0.193*** (0.05)	0.124*** (0.02)	-0.275*** (0.06)
Large-scale empirical articles	-0.123* (0.05)	0.005 (0.05)	-0.188*** (0.05)	0.228*** (0.03)	-0.211*** (0.06)
Theoretical articles	-0.178** (0.06)	-0.344*** (0.06)	-0.257*** (0.06)	-0.115** (0.04)	-0.395*** (0.07)
Publication year	0.013*** (0.00)	0.016*** (0.00)	0.010*** (0.00)	0.026*** (0.00)	0.007*** (0.00)
Constant	-24.767*** (2.88)	-31.598*** (2.63)	-18.489*** (3.55)	-51.724*** (2.23)	-12.217*** (3.60)
Observations	5,475	4,851	4,419	6,569	3,700
Loglikelihood	-8011	-9364	-6159	-13006	-5290
Mc Fadden's Pseudo R ²	0.00628	0.0109	0.00484	0.0259	0.00453
Reject H0: $\alpha=0$	$p > .05$	$p < .001$	$p > .05$	$p < .001$	$p > .05$
Standard errors in parentheses					
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$					

Even then, the content type “large-scale/comparative” has a significant, negative impact on the number of authors per article compared to the content type “experimental” in Economics, Political Science, and Sociology, meaning that large-scale/comparative articles have significantly lower author numbers than experimental articles. This effect is reversed for Psychology: The content type “large-scale/comparative” has a positive effect on the number of authors per article, compared to the content type “experimental”. H3d can be accepted.

CONCLUSIONS

The current findings provide confirmation that the average number of authors is on the rise in all analysed fields of the Social Sciences. However, different authorship patterns were discovered: Developments in Psychology and Educational Sciences are more in line with trends in STEM fields such as high shares of co-authored papers and multi-authorship as the most common form of publication. In contrast, single authorship is still dominant in Political Science and Sociology.

All three analysed factors, publication year, international collaborations, and article content were found to be significantly correlated with the number of authors. As the last two are confounded with time itself, as IC are on the rise, and more articles are published now can be

⁶ Because of over-dispersion, NBReg has to be used for Educational Sciences and Psychology.

classified as experimental/comparative, publication year was also controlled for. Still, both factors are significantly correlated with author numbers in all analysed disciplines.

Differences between the disciplines were found for article content and the number of authors; Experimental articles have on average more authors than large-scale empirical articles in Economics, Political Science and Sociology, whereas large-scale articles have on average more authors in Psychology.

This study's significant findings show that aggregating disciplines from the Social Sciences may not always be appropriate, as differences in authorship patterns exist that could not be observed without separate analyses. In addition, articles' content types were found to be significantly correlated with author numbers, so this factor should be considered for further analyses.

However, this analysis is restricted to five journals with the highest impact factor in each discipline. It would further strengthen the results to repeat this study with more journals, or probably with a set with journals from the top- and the bottom-quantile. Furthermore, this article focusses on English-language journals from the United States/Western Europe. Analysing non-English journals or journals from other parts of the world, would allow the introduction of a cross-cultural perspective on authorship patterns in the Social Sciences.

REFERENCES

- Chompalov, I. (2014). Lessons Learned from the Study of Multi-organizational Collaborations in Science and Implications for the Role of the University in the 21st Century. In M. Herbst (Ed.), *The Institution of Science and the Science of Institutions: The Legacy of Joseph Ben-David* (pp. 167–184). Dordrecht: Springer Netherlands.
- Cronin, B., Snyder, H., & Atkins, H. (1997). Comparative citation rankings of authors in monographic and journal literature: A study of sociology. *Journal of Documentation*, 53(3), 263–273.
- Cunningham, S. J., & Dillon, S. M. (1997). Authorship patterns in information systems. *Scientometrics*, 39(1), 19–27.
- Eifler, S. (2014). Experiment. In N. Baur & J. Blasius (Eds.), *Handbuch Methoden der empirischen Sozialforschung* (pp. 195–209). Wiesbaden: Springer Fachmedien Wiesbaden.
- Endersby, J. W. (1996). Collaborative Research in the Social Sciences: Multiple Authorship and Publication Credit. *Social Science Quarterly*, 77(2), 375–392.
- Engels, T. C. E., Ossenblok, T. L. B., & Spruyt, E. H. J. (2012). Changing publication patterns in the Social Sciences and Humanities, 2000–2009. *Scientometrics*, 93(2), 373–390.
- Gerber, A. S., & Green, D. P. (2013). Field Experiments and Natural Experiments. In R. E. Goodin (Ed.), *The Oxford Handbook of Political Science* (pp. 357–384). Oxford University Press.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The new production of knowledge: The dynamics of science and research in contemporary societies* (Reprinted.). Los Angeles: Sage Publ.
- Glänzel, W. (2002). Co-authorship patterns and trends in the sciences (1980-1998). A bibliometric study with implications for database indexing and search strategies. *Library Trends*, 50, 461–473.
- Glänzel, W., & Schubert, A. (2004). Analysing Scientific Networks Through Co-Authorship. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of quantitative science*

- and technology research. *The use of publication and patent statistics in studies of S&T systems* (pp. 257–276). Dordrecht: Kluwer Acad. Publ.
- Gorard, S. (2004). *Quantitative methods in Social Science* (reprinted.). New York, N.Y.: Continuum.
- Hattke, F., Blaschke, S., & Frost, J. (2014). Governance Logics in Universities. In A. M. Pettigrew, E. Cornuel, & U. Hommel (Eds.), *The Institutional Development of Business Schools* (pp. 69–94). Oxford University Press.
- Hicks, D. (1999). The difficulty of achieving full coverage of international Social Science literature and the bibliometric consequences. *Scientometrics*, 44(2), 193–215.
- Hoekman, J., Frenken, K., & Tijssen, R. J. (2010). Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy*, 39(5), 662–673.
- Huang, D.-w. (2015). Temporal evolution of multi-author papers in basic sciences from 1960 to 2010. *Scientometrics*, 105(3), 2137–2147.
- Hudson, J. (1996). Trends in Multi-Authored Papers in Economics. *The Journal of Economic Perspectives*, 10(3), 153–158.
- Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26(1), 1–18.
- Kliegl, R., & Bates, D. (2011). International collaboration in psychology is on the rise. *Scientometrics*, 87(1), 149–158.
- Levsky, M. E., Rosin, A., Coon, T. P., Enslow, W. L., & Miller, M. A. (2007). A descriptive analysis of authorship within medical journals, 1995-2005. *Southern medical journal*, 100(4), 371–375.
- Leydesdorff, L., Park, H. W., & Wagner, C. (2014). International coauthorship relations in the Social Sciences Citation Index: Is internationalization leading the Network? *Journal of the Association for Information Science & Technology*, 65(10), 2111–2126.
- Martin, M. W., & Sell, J. (1979). The Role of the Experiment in the Social Sciences. *The Sociological Quarterly*, 20(4), 581–590.
- Nederhof, A. J. (2006). Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A Review. *Scientometrics*, 66(1), 81–100.
- Nowotny, H., Scott, P., & Gibbons, M. (2001). *Re-Thinking Science: Knowledge and the Public in an Age of Uncertainty* (1. Aufl.). s.l.: Polity.
- Sin, S.-C. J. (2011). International coauthorship and citation impact: A bibliometric study of six LIS journals, 1980–2008. *Journal of the American Society for Information Science and Technology*, 62(9), 1770–1783.
- Slaughter, S., & Leslie, L. L. (1997). *Academic capitalism: Politics, policies, and the entrepreneurial university*. Baltimore: Johns Hopkins University Press.



ERIH PLUS – Making the SSH visible, searchable and available

Gry Ane Vikanes Lavik* and Gunnar Sivertsen**

* gry.lavik@nsd.no

Norwegian Centre for Research Data (NSD)
Harald Hårfagres gate 29, N-5007 Bergen, Norway

** gunnar.sivertsen@nifu.no

Nordic Institute for Studies in Innovation, Research and Education (NIFU)
P.O. Box 2815 Tøyen, N-0608 Oslo, Norway

ABSTRACT

The European Reference Index for the Humanities and the Social Sciences (ERIH PLUS) is a non-commercial initiative to increase the visibility of the humanities and the social sciences (SSH) by providing an international infrastructure for a comprehensive bibliographic coverage of the scholarly communication and publishing in these fields. A well-defined, standardized and dynamic register of scholarly journals and series is already up and running. Since a combination of national and international publishing is practised in the SSH, titles in all European languages are covered in ERIH PLUS if they correspond to six verifiable criteria. The register thereby challenges the commercial indexing services which tend to represent the SSH as “peripheral” to the sciences.

Keywords

Scholarly journals; journal list; social sciences; humanities; bibliometrics; research evaluation; current research information systems.

INTRODUCTION

The general aim of the European Reference Index for the Humanities and the Social Sciences (ERIH PLUS) project, which was originally initiated by the European Science Foundation, is to increase the visibility of the humanities and the social sciences (SSH) by providing an international infrastructure for a comprehensive bibliographic coverage of the scholarly communication and publishing in the fields. The register of scholarly journals and series presently covers around 7,500 journals and series in the SSH. It is dynamic and open to new journals all the time. Titles and identifiers are updated continuously against the international ISSN register. The ERIH PLUS project has been implemented and is further developed at:

<https://dbh.nsd.uib.no/publiseringskanaler/erihplus/about/index>

The relevance of the ERIH PLUS project for STI ENID in 2016 is related to the focus on the SSH and the understanding of peripheries as areas that are not adequately covered or targeted by current indicators. ERIH PLUS is one of several responses to the need for developing approaches and indicators that provide a more accurate or valid representation of scholarly publishing and communication in the SSH. It is one of the conditions for establishing more adequate bibliometric data in the SSH (Martin et al. 2010), particularly if the general solution is an integration of national current research information systems (Hicks & Wang 2009). This

paper will be the first international presentation of the ERIH PLUS project at an STI conference.

ERIH PLUS provides research information and evaluation systems with a well-defined, standardized and dynamic register of scholarly journals and series in the SSH. An example of a possible application of ERIH PLUS is to standardize bibliographic data and make them available and comparable across different current research information systems (CRIS) that are used for statistics or evaluation. An example of this type of application is the recommendation from the Technopolis Group to the Czech government in 2015 on the further development of the Czech national CRIS system *RD&I IS*. Arnold & Mahieu (2015) suggest:

enhancing the value of the RD&I IS for evaluation through the development of a standardised and dynamic register of scholarly journals, series and book publishers, similar to the ones developed in countries like Belgium (the Flanders) and Norway. This would be especially useful for SSH, combining for example the current List of peer-reviewed non-impact journals published in the Czech Republic with ERIH PLUS.

ERIH PLUS is a deliberate attempt to go beyond the commercial indexing services such as Web of Science and Scopus by covering more comprehensively all peer-reviewed scholarly journals in the SSH that are publishing at a minimum national level. Even with the latest expansions of the commercial databases, there is need for acknowledging journals of good quality in the SSH that are not yet covered (Sivertsen, 2014).

BACKGROUND: FROM ERIH TO ERIH PLUS

ERIH (the European Reference Index for the Humanities) was originally created and developed by European researchers under the coordination of the Standing Committee for the Humanities (SCH) of the European Science Foundation (ESF). The ERIH lists, which initially mainly covered disciplines in the humanities, were first published by ESF in 2008, while revised lists were made available in 2011-2012.

In 2014, the responsibility for the maintenance and operation of ERIH was transferred to the Norwegian Centre for Research Data (NSD), a non-commercial organization owned by the Norwegian Ministry of Education and Research. NSD also runs the Norwegian Register of Scientific Journals, Series and Publishers as a resource for *CRISStin* (the Current Research Information System in Norway).

The international register of journals and series at NSD is now called *ERIH PLUS* in order to indicate that it has been extended to *include the social sciences*. Two other changes, which will be explained below, have been made in collaboration with the SCH of the ESF:

1. Journals are not ranked any more (A, B, C). Instead, more objective criteria for inclusion have been established.
2. The register has become dynamic. Instead of having expert panels meet at large intervals, NSD is responsible for a daily operation and development of the register with the aid of National Experts and an international Advisory Group.

The classification of journals in categories A, B and C, all though used in some countries for the allocation of institutional funding, was one of the most controversial elements of the

initial ERIH lists. The controversy rose from the notion that the inclusion and ranking of journals in some cases seemed arbitrary (KNAW 2011). Another side that made the list open to criticism was that the ranking of journals was done within each discipline, which made journals operating across disciplines difficult to rank. More controversial though, was the fact that the ranking seemed to strengthen the focus and pressure towards publishing in the English-speaking world, while weakening publications in national languages. When the attention then was turned towards strengthening publications within national language groups, yet another concern was voiced, namely that there were cases of national journals of a poor quality that were ranked too high on the basis of a recommendation coming from only one committee member (KNAW 2011). All in all the criticism against the initial lists were in some way or another connected to the system of ranking the journals.

To meet the criticism and to further develop ERIH PLUS in a more useful and maybe fruitful direction, several factors were changed when transferring ERIH to NSD. The criteria for inclusion, the review process and the categorization are now handled differently as ERIH PLUS is provided with a continuous inclusion of new journals.

CRITERIA FOR INCLUSION

The approval procedures of the ERIH PLUS database are, as mentioned, different from the original ESF procedures. There is no peer review by expert panels. Instead, all journal submissions are treated in a standardized way and reviewed for compliance with more objective criteria that can be checked against evidence. The new criteria have been developed jointly by ESF and NSD:

1. Explicit procedures for external peer review.
2. Academic editorial board, with members affiliated with universities or other independent research organizations.
3. Valid ISSN code, confirmed by the international ISSN register.
4. All original articles should be accompanied by abstracts in English and/or another international language relevant for the field.
5. Information about the affiliations and addresses of the authors should be published for each article.
6. Minimum national level: No more than two thirds of the authors published in the journal are from the same institution.

Criteria 1, 2, and 6, taken together, ensure that ERIH PLUS promotes research quality in the SSH. By allowing for journals published in the national languages, societal relevance is also promoted. Criteria 3-5, taken together, ensure that data will be efficiently relevant, searchable and comparable across Current Research Information Systems and other bibliographic data sources. They are also required for performing bibliometric analysis with the use of data from Current Research Information Systems (Sivertsen 2016b).

All six criteria must be assessed as fulfilled in order for a journal to be included in ERIH PLUS. When assessing that the journal complies with the criteria, only publicly available information is reviewed. It will not suffice that a scholar affiliated with the journal can give a description of how for example the peer review is conducted. If the peer review procedures are not published on the journals' web site or in the printed version, the criterion concerning

peer review is regarded as not being fulfilled. Transparency is central in the ERIH PLUS project.

Transparency is believed to ensure truthfulness (Habermas 1962). The ERIH PLUS transparency-requirement, must not however, be confused with requiring the journals to be open access (although ERIH PLUS can promote open access, see below). It is merely the information indicating that a journal complies with the criteria that has to be assessable, and not only for reviewers, but for all. The easiest way of enabling assessment of the criteria would be to demand that all information about the criteria must be available on-line. And in the first months of ERIH PLUS, this was indeed a requirement. However, while all journals do have a web site, it has become evident that far from all journals in the SSH have all information about for example authorship of the articles on-line. Therefore, when processing journals' applications for inclusion, on-line transparency is currently only advised for, not demanded. Journals are included also if documentation showing information about the criteria is given merely in the printed publication. Reviewing thousands of SSH journals have shown that on-line transparency has not yet fully dawned on the world of SSH publishing, all though the SSH seems to be moving in that direction. This point we shall return to in the section "Preliminary effects and achievements" below.

PROJECT ORGANIZATION

ERIH PLUS has a web site for on-line submissions at NSD, where journals are submitted continuously. The submissions may only come from scholars affiliated with universities and other independent research organizations. Journal editors, librarians and members of scholarly associations may submit if they meet this general requirement. Submissions from commercial publishers will not be considered. When submitting a journal, the scholar has to give information about title, ISSN, URL, language and publisher for the journal to be further evaluated.

Each individual journal is reviewed in a standardized way by reviewers looking for formal criteria. If needed, advice is sought from appointed National Experts. Being first and foremost a European index, ERIH PLUS has a National Expert representing every European country with more than 10 journals submitted for inclusion in the index. The experts are scholars who most often are affiliated with a university, and they have knowledge about and interest in scholarly publishing. The National Experts give guidance and advice in cases where the NSD reviewers are in doubt about a journal or a publisher. In addition to the National Experts, an international Advisory Group, also consisting of scholars, gives guidance about strategy and principal matters. The group represents different countries and fields in the SSH and has founders of the original ERIH project among its members.

When journals are reviewed by professional reviewers who can seek advice from appointed National Experts, this provides the project with control by the scholarly communities, while at the same time enabling a continuous flow of submissions and new inclusions to the register. The project organization is thus believed to be time and cost efficient and the standardization of the process might prevent cases of arbitrariness that the ERIH-list was criticized for as referred to above.

The ERIH PLUS web site is, furthermore, a place to find information about approved journals. A freshly updated list of all included journals with ISSN, country of publication and discipline can be downloaded at all times.

PRELIMINARY EFFECTS AND ACHIEVEMENTS

One effect of ERIH PLUS becoming a dynamic register with evidence-based criteria is as suggested above an increasing awareness about on-line transparency among researchers involved in SSH journals throughout Europe. This is particularly evident from the communication with the about 52 per cent of the submitted journals that were *not* approved after the first submission. In many of these cases, the journals actually did fulfil the criteria, but the information had not been made publicly available or was available only in the print version. Many of the journal editors have now decided to change their practice towards greater transparency about the journal's organization and content.

Those journals which were refused because they only had information about the editorial board, the peer review procedure and the authors of the articles in the printed version, have in many cases adopted a practice of indicating this information on the web site of the journal as well. These journals have been re-submitted for evaluation and approved for inclusion in ERIH PLUS. Other journals might have had information about editorial board and the peer review procedure as internal documents only. After being informed that the refusal was a result of their lack of practicing transparency, many of these journals are now practising on-line transparency. Based on this experience we can for a fact say that many European journals within SSH are now giving more detailed information about the editorial board, the peer review procedures, and about authors as well as abstracts in English on their web site, all in accordance with the ERIH PLUS criteria. Journal editors and staff have thus helped ERIH PLUS with the aims of making scholarly publications in the SSH more visible, searchable and available across Europe.

FUTURE AIMS FOR ERIH PLUS

The aims for ERIH PLUS have been set as:

- Continue to challenge the limited coverage of the SSH in the commercial data sources.
- Become increasingly important as a resource for the running of CRIS-systems worldwide.
- Guide a more proper coverage of the SSH in performance-based funding systems.
- Become an agreed standard and resource in the development of a proper evaluation.
- Become an important information source for supporting research quality in OA journals.
- Develop from being a "list" to becoming an enriched information resource - through European collaboration.

A background for these aims can be found in a report from a working group appointed by the European Science Foundation to discuss the future of ERIH in the wider perspective of international research communication. The working group decided to include the social sciences in the perspective. Then they compared the SSH to science, technology and medicine and asked how the scholarly literature in the SSH could be made searchable and available across countries and languages in the same way. A link to the report is given in the references below.

REFERENCES

More detailed information about ERIH PLUS is given at:

<https://dbh.nsd.uib.no/publiseringskanaler/erihplus/about/index>

The report of the ESF working group in 2010, *Towards comprehensive bibliographic coverage of the scholarly literatures in the humanities and social sciences*, is available at:

https://dbh.nsd.uib.no/publiseringskanaler/resources/pdf/ERIH_Report_from_a_working_group.pdf

Royal Netherlands Academy of Arts and Sciences (KNAW): Quality Indicators for Research in the Humanities (2011) Available at: <http://www.know.nl/en/news/publications/quality-indicators-for-research-in-the-humanities>

Arnold, E. & Mahieu, B. (2015). *R&D Evaluation Methodology and Funding Principles. Summary Report*. Brighton: Technopolis Group.

Hicks, D. and Wang, J. (2009). *Towards a Bibliometric Database for the Social Sciences and Humanities. A European Scoping Project* (Appendix 1 to Martin et al., 2010). Arizona: School of Public Policy, Georgia University of Technology.

Martin, B., Tang, P., Morgan, M., Glänzel, W., Hornbostel, S., Lauer, G., et al. (2010). *Towards a bibliometric database for the social sciences and humanities—A European scoping project: A report produced for DFG, ESRC, AHRC, NWO, ANR and ESF*. Sussex: Science and Technology Policy Research Unit.

Sivertsen, G. (2014). Scholarly publication patterns in the social sciences and humanities and their coverage in Scopus and Web of Science. In E. Noyons (Ed.), *Proceedings of the science and technology indicators conference 2014 Leiden* (pp. 598-604). Leiden: CWTS.

Sivertsen, G. (2016a). Patterns of internationalization and criteria for research assessment in the social sciences and humanities. *Scientometrics*, 107(2), 357-368

Sivertsen, G. (2016b). Data integration in Scandinavia. *Scientometrics*, 106, 849-855.

Habermas, J. (1962) *Strukturwandel der Öffentlichkeit: Untersuchungen zu einer Kategorie der bürgerlichen Gesellschaft*. Neuwied, Berlin: Luchterhand.

Martin, B., Tang, P., Morgan, M., Glänzel, W., Hornbostel, S., Lauer, G., et al. (2010). *Towards a bibliometric database for the social sciences and humanities—A European scoping project: A report produced for DFG, ESRC, AHRC, NWO, ANR and ESF*. Sussex: Science and Technology Policy Research Unit.



Indexed University presses: overlap and geographical distribution in five book assessment databases¹

Jorge Mañana-Rodríguez*, Elea Giménez-Toledo**

*jorge.mannana@cchs.csic.es

**elea.gimenez@cchs.csic.es

Centro de Ciencias Humanas y Sociales, (ÍLIA Research Group), CSIC, Albasanz Street, 28037, Madrid (Spain)

ABSTRACT:

Scholarly books have been a periphery among the objects of study of bibliometrics until recent developments provided tools for assessment purposes. Among scholarly book publishers, University Presses (UPs hereinafter), subject to specific ends and constrains in their publishing activity, might also remain on a second-level periphery despite their relevance as scholarly book publishers. In this study the authors analyze the absolute and relative presence, overlap and uniquely-indexed cases of 503 UPs by country, among five assessment-oriented databases containing data on scholarly book publishers: Book Citation Index, Scopus, Scholarly Publishers Indicators (Spain), the lists of publishers from the Norwegian System (CRISTIN) and the lists of publishers from the Finnish System (JUFO). The comparison between commercial databases and public, national databases points towards a differential pattern: prestigious UPs in the English Speaking world represent larger shares and there is a higher overall percentage of UPs in the commercial databases, while the richness and diversity is higher in the case of national databases. Explicit or *de facto* biases towards production in English by commercial databases, as well as diverse indexation criteria might explain the differences observed. The analysis of the presence of UPs in different numbers of databases by country also provides a general picture of the average degree of diffusion of UPs among information systems. The analysis of ‘endemic’ UPs, those indexed only in one of the five databases points out to strongly different compositions of UPs in commercial and non-commercial databases. A combination of commercial and non commercial databases seems to be the optimal option for assessment purposes while the validity and desirability of the ongoing debate on the role of UPs can be also concluded.

Key Words: University Presses, Scholarly Books, scholarly book assessment, database coverage, research evaluation.

¹ This research has been carried out in the framework of the Project CSO2015-63693, ‘Las prensas universitarias iberoamericanas y el libro científico en español: calidad, modelos de negocio y política institucional’ funded by the Spanish Ministry of Economy and Competitiveness. The authors gratefully acknowledge Gunnar Sivertsen and Janne Pölonen for the information provided regarding book publishers’ lists.

INTRODUCTION

Books have remained on the peripheries of bibliometrics for a long time. Despite the emergence of bibliometric information systems in the 70's, it was not until 1996 that Eugene Garfield proposed the creation of a database for scholarly books (Garfield, 1996). Books were and are a key communication channel for Social Scientists and Humanists (Hicks, D., 2004; Engels, Ossenblok & Spruyt, 2012; Thompson, 2002; Giménez-Toledo *et al.*, 2016). 53% of the output in SSH fields in Norway were published, between 2005 and 2009, in the form of monographs and book chapters (Sivertsen & Larsen, 2012) while the percentages in Finland were (for 2011-2012) 39 % and 47% in the case of Social Sciences and Humanities respectively (Puuska, 2014); 62% of the output of Spanish universities in Arts and Humanities were Books and Chapters (Michavila, 2012). Also, citation analysis considering books as a source of reference information shows that its relevance is far from residual (Gorraiz *et al.*, 2013; Leydesdorff & Felt, 2012). Finally, in market terms, scholarly books represent a large percentage of total profits in the book market in Europe: 19.5%, being 'the second most important sales segment, after consumer (trade) books', in 2014 (Federation of European Publishers, 2014); in the case of Spain, 3.8% of the yearly turnover of the book industry corresponds to that from scientific-technical books and 10.8% to books in fields of the Social Sciences and the Humanities (FGEE, 2014).

University Presses, as part of the scholarly book publishing sector, are not necessarily a periphery in the scholarly book segment: 11,000 books a year are published by the ninety-two university presses belonging to the American Association of University Presses; Abel & Newlin, 2002, while 10% of all book publishing in Latin America was produced, in 2013, by University Presses (CERLALC, 2014, p. 30). In the case of Spain, University Presses published 6.5% of the total volume of books published in Spain in 2013 (FGEE, 2014) and the Presses Universitaires de France keep a catalogue with over 5,000 titles (PUF). Despite not being in the periphery in terms of publication volume and share, they can be considered a periphery since books are mostly produced in SSH disciplines and SSH is a small part of research in terms of funding and human resources. Moreover university presses are not, in most cases, privately held companies but do compete with private companies while keeping a set of specificities (AEUP, 2015) that might set them closer to the periphery than to the core of scholarly book publishers: University Presses are often constrained by normative obligations from the entities they belong to. Often characterized by a local factor with regard to the works published (AAUP, s.d.), the languages used and the specific factors which regulate their publishing activity, it is the aim of this work to analyze the role of University Presses in five assessment oriented databases: Book Citation Index, Scopus, the lists from the CRISTIN system in Norway, the Finnish Lists and Scholarly Publishers Indicators. It is assumed that University Presses tend to be closely related to the activity of the university and, therefore the diffusion among different information systems shows recognition far from the closest institution or region. That international presence can be related to variables such as: a) recognition of the publisher by foreign specialists, b) improvements in the diffusion strategies of the publishers, c) professionalization and budget in marketing tasks, d) publication languages, e) business model and f) topics covered by the publisher (local topics would be less interesting for audiences abroad). The first step in the study of those conditioning variables is the analysis of the currently available data. Since the variables which explain the indexation of a given publisher differ, the comparisons should be done taking it into

consideration. There can be identified three systems for the inclusion of book publishers in the different information systems:

- a) Book Citation index and Scopus apply different criteria including reputation and impact or content quality and, in the case of BCI: “English language full text is highly desirable, but books with full text in a language other than English are also considered for coverage in Book Citation Index”²
- b) Finnish lists and Norwegian lists include scholarly publishers in which scholars from the respective countries have published research.
- c) SPI includes book publishers which have been mentioned as relevant by a set of Spanish scholars through a survey methodology.

OBJECTIVES

The five databases studied here have in common the fact that they are recognition-based systems. The objectives of this work are the following ones:

- a) Identify descriptive patterns in the geographical distribution of the UP’s in the five databases.
- b) Compare the coverage of the privately held databases to the coverage of the public databases
- c) Identify the degree of overlap between the different databases.
- d) Extract conclusions on the applicability of the different systems for assessment at the national or international level.
- e) Identify variations in the visibility or recognition of UPs throughout the databases and which role they play in each one.

METHODOLOGY

Data origin:

The origin of data can be traced to the development Scholarly Publishers Indicators Expanded: http://ilia.cchs.csic.es/SPI/expanded_index.html), which included the five main databases on scholarly books. The lists of publishers were retrieved between December, 2015 and February, 2016 from the following sources:

-SPI: <http://ilia.cchs.csic.es/SPI/>

-Book Citation Index: <http://wokinfo.com/mbl/publishers/>

-Scopus:

http://www.elsevier.com/_data/assets/excel_doc/0005/154571/Scopus_books_29_4_15.xls
[X](#)

-Norwegian lists (CRISTIN): institutional exchange of files /Personal communication with Gunnar Sivertsen.

-Finnish lists: <http://www.tsv.fi/julkaisufoorumi/lataa.php?id=>

Although the selected sources allow multiple analysis, to be focused on UP let us know more about the role and the behaviour of one of the peripheries (UP) of the peripheries (evaluation of SSH)

Data processing: First, a master list containing the names of the publishers in all the five information systems was prepared; then, cleansing operations were performed: deletion of

² http://wokinfo.com/media/pdf/BKCI-SelectionEssay_web.pdf

non ASCII symbols, deletion of spaces, etc.,. Then, an exact match search was performed with the master list and each individual list. The second phase involved the manual normalization of variants. The process yielded a total of 3948 distinct book publishers. The third phase of data processing involved the identification of university presses. For this purpose, the following word roots were searched in the names of the book publishers (Table 1):

Table 1: character chains used for the automatic identification of Ups.

Univers	Root of the term “university” in most romance languages.
Yliop*	Root of the term “Yliopisto”, University in Finnish.
Korkea*	Root of the term “korkeakoulu”, College and / or synonym of University in Finnish
Colleg	Root of the term “College” in most romance languages
Ülikoo	Root of the word ‘Ülikool’, University in Estonian
Egyet / ösz	Roots of the word “Egyetem” in Hungarian
Uniwers	Root of the word Uniwersytet in Polish

* Janne Pölönen, Coordinator of the Publication Forum, Federation of Finnish Learned Societies, was contacted by the researchers regarding the issue: he provided a wider list of Finnish UP’s not necessary containing the roots detailed in this table as well as useful clarifications on the scheme for inclusion of UPs in the Finnish system.

Some other roots were also tested (I.E. for Pinyin Chinese and Hepburn Rōmaji Japanese “Daigak”) with limited results.

Once the university presses were identified, a final manual depuration was carried out, excluding erroneously identified cases as well as normalizing variants. Also, a further review of the full list of publisher was manually carried out in order to identify university presses not previously identified.

Once the set of UPs was identified, the number of different databases in which each UP was included was counted. From this point, descriptive statistics were computed.

RESULTS

Figure 1. Number of book publishers in each database (total n=3948)



The largest set of book publishers can be found in the Finnish System, followed by the Norwegian Lists and SPI (Fig. 1).

Figure 2. Distribution of UPs among databases

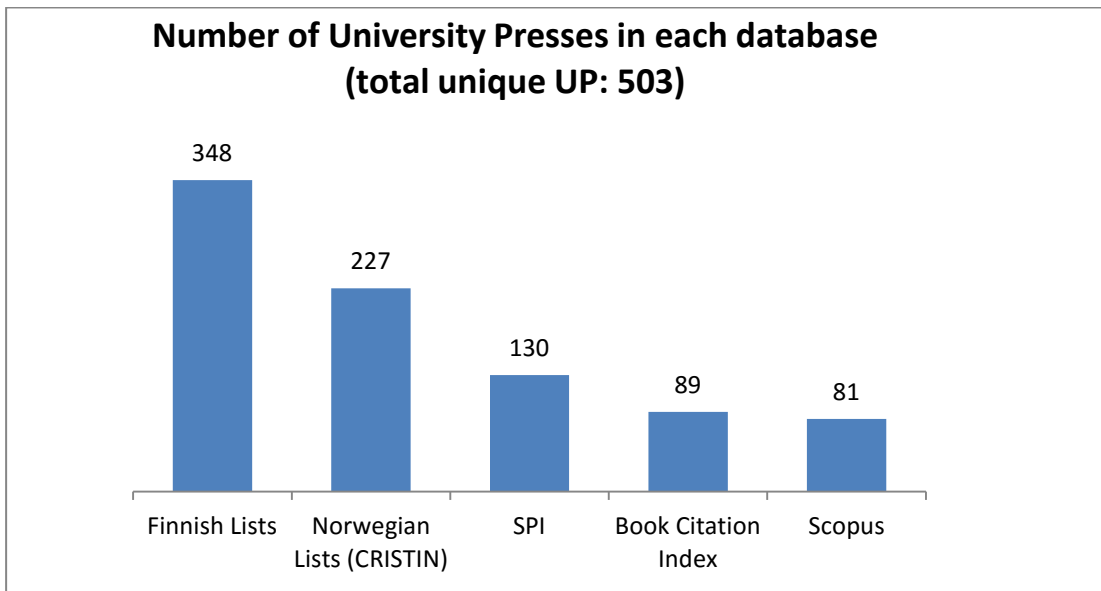
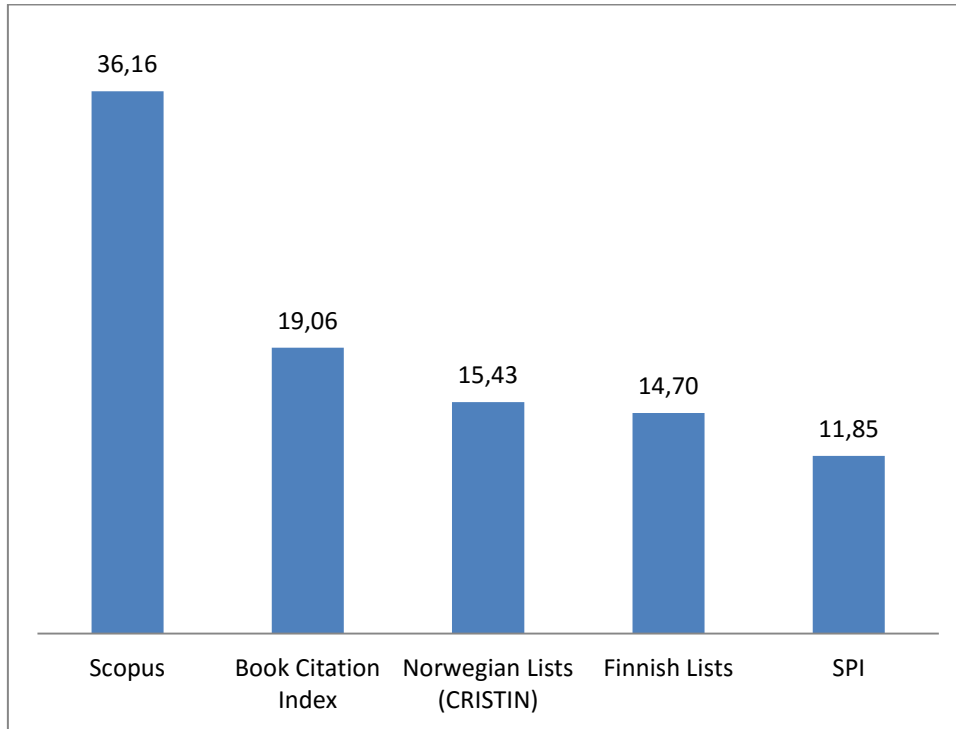
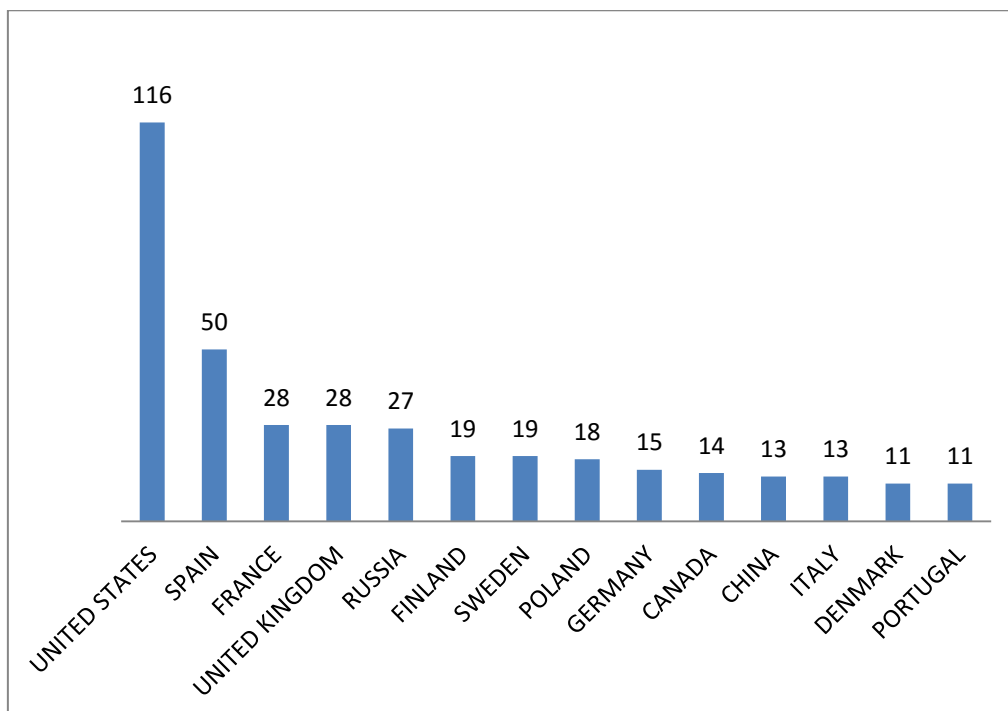


Figure 3. Percentage of UPs in each database



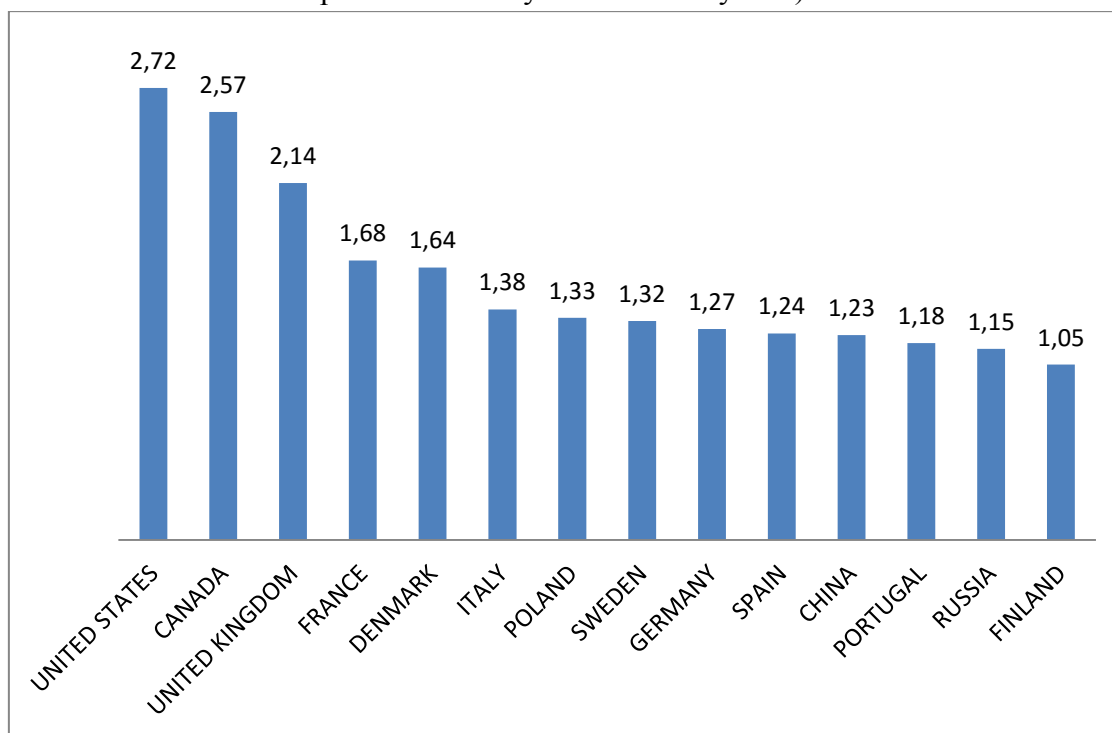
The percentage of UPs is significantly higher in the commercial databases, remarkably in Scopus, while SPI remains the database with the lowest percentage of University Presses (Fig. 3).

Figure 4. Countries with at least 10 UPs in any information system



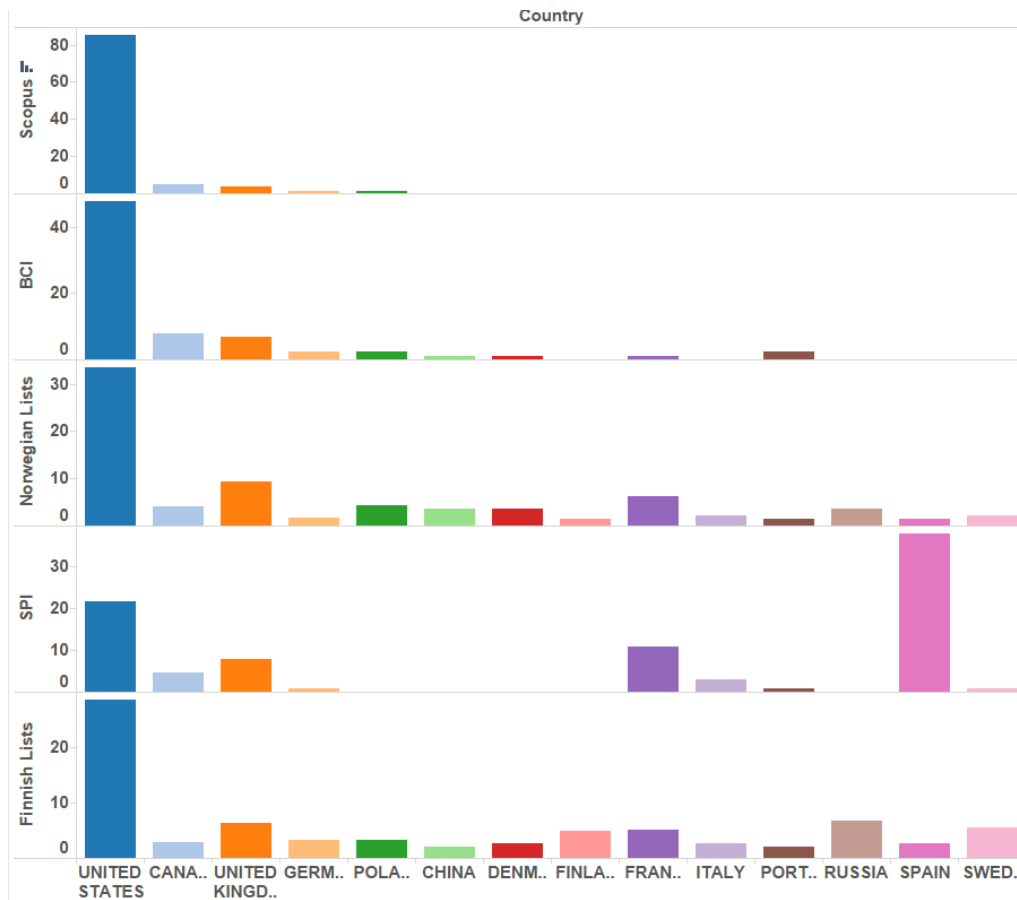
USA, Spain, France and Russia are the countries with a higher presence of UPs in any of the five systems (Fig. 4).

Figure 5. Average number of information systems for each country's UPs (if 10 or more book publishers in any information system).



The United States, Canada and the UK are the countries which UPs show the higher average presence in the five information systems (Fig. 5).

Figure 6. Distribution of each country UPs among databases



A thematic map with this information can be found at: https://public.tableau.com/shared/ZZZGY8DXS?:display_count=yes

A large diversity of different countries in the public sources of Spain, Norway and Finland can be observed, while the concentration of countries (in most cases English-Speaking ones) in the case of the privately held products is observably larger.

COMBINATIONS, OVERLAP & ‘ENDEMIC’ UPS

The frequency of combinations in the five databases shows great variability in the overlap pattern, but also that a large number of UPs are indexed only in one database. Using a parallelism with the term used in biology we decided to term those UPs as ‘endemic’. The full set of combinations showing the degree of overlap between the different databases can be found at:

<https://public.tableau.com/profile/publish/Frequencyofoverlapping/Sheet1#!/publish-confirm>

Table 2. ‘Endemic’ University Presses

Database	Number of ‘Endemic UPs’	% respect total n of UPs	Composition (country of the UPs with higher frequency among ‘endemic’ Ups)	
			Country	Frequency
Finnish Lists	146	42,2	Russia	19
SPI	65	50,0	Spain	39
Norwegian Lists (CRISTIN)	48	21,1	China, Poland and United Kindgom	5
Book Citation Index	24	27,0	Czech Republic	5
Scopus	5	6,2	United States	5

The large percentages of endemic UPs in the Finnish Lists and SPI strongly contrasts with the very low percentage of endemic UPs in the case of Scopus, while the lack of coincidence in the composition section of the table point towards a distinctive composition the ‘endemisms’ (Table 2).

CONCLUSIONS

Databases produced at the national level enable a clearer representation of the richness and diversity of book publishers relevant to the scholarly publishing activity of each country. National book publishers are better known, their catalogues are closer to the authors and those publishers might be more accessible for certain topics. Commercial databases might tend to be highly selective, thus possibly more prone to choose the least particular publishers of each country, region or language.

Also, some publishers with international scope and reach which might not be well known by experts in some countries. The direct association between presence in a given database and intrinsic quality should be avoided, being a combination of national and international classifications the best option for evaluation purposes.

The analysis of these data shows that there might be a close relation between the topics published and the readers of the works: local issues might play a significant role in national level book publishers, which does not necessarily imply a lower quality of the work. Nevertheless, the difference in composition and coverage and the analysis of the overlap of the various information systems shows that what the public and private databases cover is largely different.

The limited number of publishers covered by BCI and Scopus (both commercial databases) evidences their restrictiveness for assessment purposes. Considering only the coverage towards UPs, commercial databases show a larger percentage, while in the case of the Spanish and Finnish systems the percentages are particularly low. Taking into account how the latter databases have been constructed, it can be concluded that scholars in these countries consider UPs less prestigious (Spain) or choose to publish less (Finland) in those publishers.

USA, UK and Canada UPs show high averages of presence among the different information systems (Figure 5): considering the bias towards English-publishing publishers in the commercial databases, UPs might occupy a similar position, in terms of recognition to publishers. Greater diversity can be observed among the three public databases concerning the number of countries with publishers in each database. Political and geographical influence

could be a factor contributing to this observation: Russian, Swedish, Polish and Danish UPs are covered in the Nordic countries' lists while the number of Spanish databases is low (3 in the Norwegian lists and 10 in the Finnish lists). Italy, Portugal and France are barely covered in BCI and not covered in Scopus and Finland, Sweden, Russia, Italy and Spain UPs are present in the public databases but are almost invisible for BCI or Scopus, the latter (being produced in The Netherlands) covering only two European UPs (the German Deutsche Universitäts Verlag and the Polish Deutsche Universitäts Verlag).

'Endemic' UPs could be indicative of the extent up to which local (in the nature of the topics published and / or in the languages of publication) research is relevant for each country.

Endemic' UPs in the case of Spain (SPI) are mainly Spanish while in the rest of the databases 'endemisms' are mainly from countries other than those where the headquarters of the database developers are.

DISCUSSION

The differences in the selection procedures among the five databases studied are considerable, as detailed in the introduction. Nevertheless, the composition and overlap of the databases shows a clear pattern when comparing the public databases with the private databases: in the countries where the public systems have been developed, including either publishers in which scholars have published their research or publishers which are considered the most prestigious, the diversity of publishers and the range of countries is wider than in the case of the private databases. The reasons for that observation are yet to be identified, since this study does not intend to provide a final set of explanations. Nevertheless, some ideas can be outlined. UPs indexed in all databases can be considered internationally recognized as relevant, since all the five databases are selective and imply recognition. Also, the larger diversity of publishers and countries in the public databases might correspond to the intrinsic features of research in SSH fields (often, of local interest, published in languages other than English; Hicks, 2004).

As main consequences of the analysis carried out, the sharp contrast between the coverage towards European UPs (with the exception of the UK and Germany) by the commercial and public databases can be useful in order to assess the suitability of both for assessment purposes: the latter might not be directly usable in national assessment processes. The linguistic bias of the commercial ones and the exclusion of large sets of European UPs is also a significant limitation for their use with assessment purposes for European outputs in SSH books..

The use of commercial databases which transparency is not equivalent to that of the public databases (although for understandable commercial reasons) should be the object of close scrutiny and analysis before taking a decision on its use. This is particularly relevant if the indicators provided by each database (citations in the case of the two commercial ones) are considered. Finally, since UPs are fewer than commercial databases in the public databases analyzed, the debate regarding the role of University Presses in each country, their publishing and business models might be still relevant.

REFERENCES

Abel, R. E. & Newlin, L. W. (2002). *Scholarly Publishing: Books, Journals, Publishers, and Libraries in the Twentieth Century*. US, John Wiley & Sons, Inc.

AAUP. *About university presses*. <http://www.aaupnet.org/about-aaup/about-university-presses>

AEUP (2015). <http://www.aeup.eu/aeup/about/the-need-for-aeup/>

Cerlalc/ UNESCO. 2014. *El espacio iberoamericano del libro 2014*. Available at <http://cerlalc.org/publicacion/el-espacio-iberoamericano-del-libro-2014/>

Federation of European Publishers (2014). *European Book Publishing Statistics 2014*. <http://fep-fee.eu/European-Book-Publishing-741>

Engels, T. C. E., Ossenblok, T. L. B., & Spruyt, E. H. J. (2012). Changing publication patterns in the Social Sciences and Humanities, 2000–2009. *Scientometrics*, 93, 373–390.

FGEE. 2014. *Comercio interior del libro*. Available at <http://www.mecd.gob.es/dms/mecd/cultura-mecd/areas-cultura/libro/mc/observatoriolect/redirige/estudios-e-informes/elaborados-por-el-observatoriolect/sector-libro-abril2015/sector-libro-abril2015.pdf>

Garfield, E. (1996, October). Citation indexes for retrieval and research evaluation. In: *Consensus Conference on the Theory and Practice of Research Assessment* (Vol. 7).

Giménez-Toledo, E.; Mañana-Rodríguez, J.; Engels, Tim C. E.; Ingwersen, P.; Siverten, G.; Verleysen, F. T. & Zuccala, A. A. (2016) Taking scholarly books into account: current developments in five european countries. *Scientometrics* (online first).

Gorraiz, J., Purnell, P.J., & Glänzel, W. (2013). Opportunities for and limitations of the book citation index. *JASIST*, 64(7), 1388–1398.

Hicks, D. (2004). The four literatures of social science. In: *Handbook of quantitative science and technology research* (pp. 473-496). Springer Netherlands.

Leydesdorff, L., & Felt, U. (2012). Edited volumes, monographs and book chapters in the Book Citation Index (BKCI) and Science Citation Index (SCI, SoSCI, A&HCI). *Journal of Scientometric Research*, 1(1), 28–34.

Michavila, Francisco (dir.). (2012). *La Universidad española en cifras*. Madrid: CRUE. http://www.crue.org/Publicaciones/Documents/UEC/LA_UNIVERSIDAD_ESPANOLA_EN_CIFRAS.pdf

Puuska, (2014). H.-M. *Scholarly Publishing Patterns in Finland: A comparison of disciplinary groups*. Tampere University Press.

PUF, 2015. Information available at: https://www.puf.com/Les_engagements_Puf#expertise-editeur

Sivertsen, G., and B. Larsen. (2012). Comprehensive bibliographic coverage of the social sciences and humanities in a citation index: An empirical analysis of the potential. *Scientometrics*, 91(2), 567- 575.

Thompson, J. W. (2002). The Death of the Scholarly Monograph in the Humanities? Citation Patterns in Literary Scholarship. *Libri*, 52: 121–36.



East-African Social Sciences and Humanities Publishing: A Handmade Bibliometrics Approach¹

Nora Schmidt*

*nora.schmidt@kultur.lu.se

Information Studies, Department of Arts and Cultural Sciences, University of Lund, Helgonavägen 3, Lund, 22362 (Sweden)

ABSTRACT

For Eastern Africa, very little information about the SSH knowledge production can be found from a European perspective. Adequate indicators like information-rich bibliographic databases that cover East-Africa-based journals and book publishers are lacking. This research in progress explores their indexing situation in detail, their development, which is closely connected to political history, their (non-)usage, and affiliations as well as career-stages of their authors. Furthermore, it also pays attention to East-Africa-based SSH researchers who use other publication venues. Any bibliometric analysis in this field needs to rely on manual data collection, otherwise it would be heavily biased. This study lays out the foundation for citation analyses, qualitative research on the publications' content and the self-description of East-African scholars against the background of an academic environment that is often described as “international”.

INTRODUCTION

In many East-African countries, social problems are crucial; and cultural heritage is vanishing quickly. Research funding, national as well as foreign, strongly emphasises Science, Technology, Medicine and applied research, while neglecting basic research in Social Sciences and Humanities (SSH) to large parts (Mouton, 2010). SSH facilitate the understanding of the current social, political, economic and cultural transformations of the region. For sustainable investments in research and innovation, funding decisions should be based on SSH basic research.

To follow up on the East-African intellectual discourse in SSH, especially from a place outside of this region, a proper information infrastructure needs to be in place. African research institutions, though, often face infrastructural problems starting with unstable electricity (Fari & Ocholla, 2015). Therefore, printed journals are still the key medium (Smart & Murray, 2014). Low online availability of East-African publications and content-related barriers—like subjects, structure or style—might contribute to a weak connectivity of scholarly communication with other world regions. A generally low publication output could also indicate that other ways of academic knowledge production, e.g. teaching, are at least equally important. Bibliometric methods alone cannot deliver an adequate picture.

Eastern Africa was selected as field for this study, because its academic system is rarely researched and seems to be especially small and isolated. With few exceptions, basic bibliographic data of the field can only be found in the *Quarterly Index to African Periodical*

1 I thank Fredrik Åström, Jörgen Eriksson and Wolfram Seidler for their comments that improved this paper.

Literature (QIAPL) produced by the *Library of Congress* and discontinued in 2011. By bilateral agreement, it is now included in the freely accessible bibliography *Africabib*, but cannot be reused in any central index or federated search system. For some of the more recent publications, the only larger source is *African Journals Online (AJOL)*.

BACKGROUND

The research in progress presented here is part of a larger project which aims to characterise East-African-based academic knowledge production and scholarly communication in SSH basic research against the background of supposedly international academia. First, I describe scholarly communication bibliometrically, including the analysis of corresponding citation networks. I then compare these results to the self-descriptions of this communication by researchers invited to a workshop at an East-African university. Finally, the adequacy of the label “international academia” for the social structure denoted by it has to be discussed.

RESEARCH QUESTIONS

The sub-project presented here is a bibliometric description of the East-African journals and book publishers of basic research in SSH, as well as other publishing venues of SSH authors based in this region. Three aspects are in focus:

1. *Publishing venues*: Which journals and book publishers form the East-African publishing landscape? How are they developing?
2. *Indexing*: How can bibliographic data of these publications be accessed?
3. *Authors*: Looking at affiliations and career-stages: who are the authors working with these publishers? Which East-Africa-based authors do not work with them? Which publication strategies can be deduced from that?

LITERATURE REVIEW

In the last decade, none of these questions have been approached. Nwagwu documented the bad coverage of African publications by “international” databases in 2005. The plan to establish an *African Citation Index* (Nwagwu, 2010) has not materialised to this date. For South African Science Journals, Pouris & Richter (2000) listed several reasons for publishing in local journals and Tijssen (2007) has shown that non-indexed journals have international citation impact. Although Cloete, Maasen, & Baily (2015) give a lot of relevant background information, they focus on the evaluation of individual universities, five of them in Eastern Africa, and say little about local publishing venues. Esseh (2011) investigated levels of access to literature in Sub-Saharan Africa.

METHOD

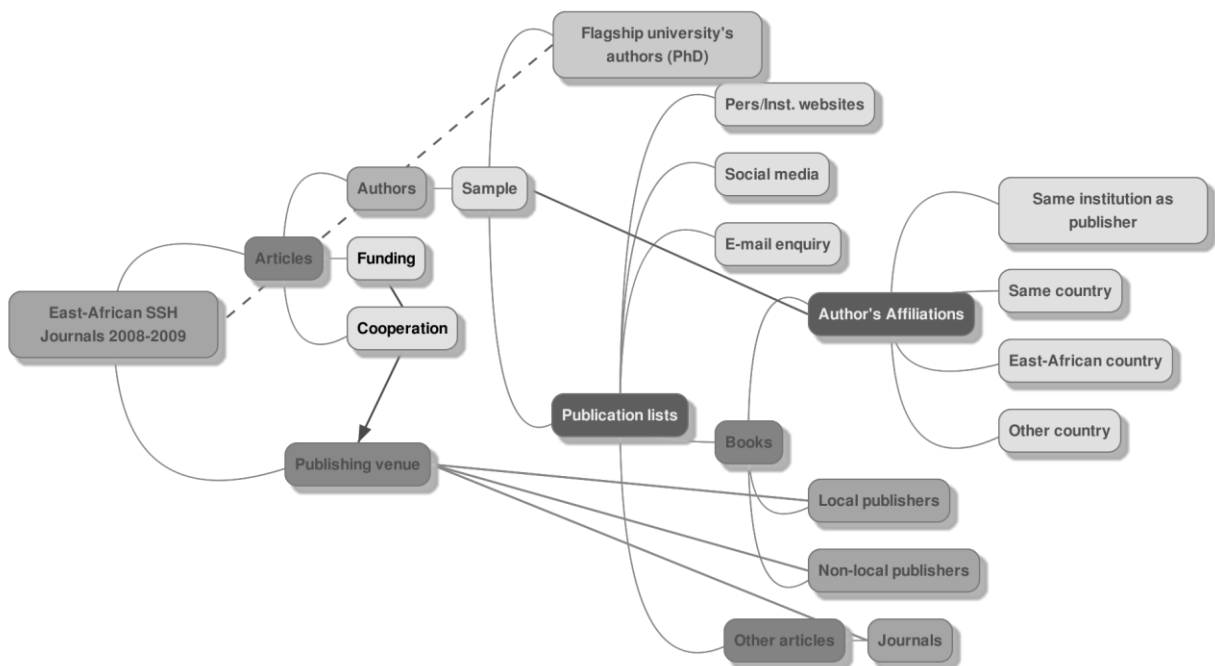
Figure 1 visualises the process of data collection and parts of the analysis. Firstly, the journals of the field, mostly print-only, have to be identified with the help of *Ulrichsweb* and *QIAPL*. To analyse citation networks, which is planned in a following part of the study on the same material basis, SSH literature has to be older than five years (Archambault & Larivière, 2010). I therefore decided to limit the data collection to the publication years 2008-2009.¹ If all preselected journals are within the scope can only be determined after looking at a sample of full texts obtained by intercontinental inter-library loan.

Unfortunately, *QIAPL* includes only basic bibliographic data, so author's affiliations have to be collected manually, resulting in a small sample. The recording of current affiliations regardless of the publishing date is sufficient, because East-African researchers are not very mobile (for Mozambique see Felleson & Mählcck, 2013). If some authors' affiliations cannot be assigned after a reasonable time spent with *Google Search*, including Social Media platforms, these articles will be removed from the sample. For all other authors, publication lists will be located, in case of need via Email request. For later analyses, CVs, information about funded projects, and collaborations will be recorded as well.

After identifying authors who use the East-African publication venues, a “control group” will suggest how many East-Africa-based researchers stay away from these: For a matching, I create a list of all SSH departments of better known universities of the region with their researchers (at least two publications), and will retrieve their publication lists etc. as well.

The publication lists reveal information about book publishing which is of high importance for SSH (see e.g. Samuels, 2013). Finally, for a sample of publication lists, I look for patterns of publication strategies until saturation is reached, paying attention to a fair distribution regarding affiliations. All publications will be categorised: 1) published with in-house infrastructures, 2) published in the same country, 3) in an East-African country, or 4) abroad.

Figure 1. Design of the study



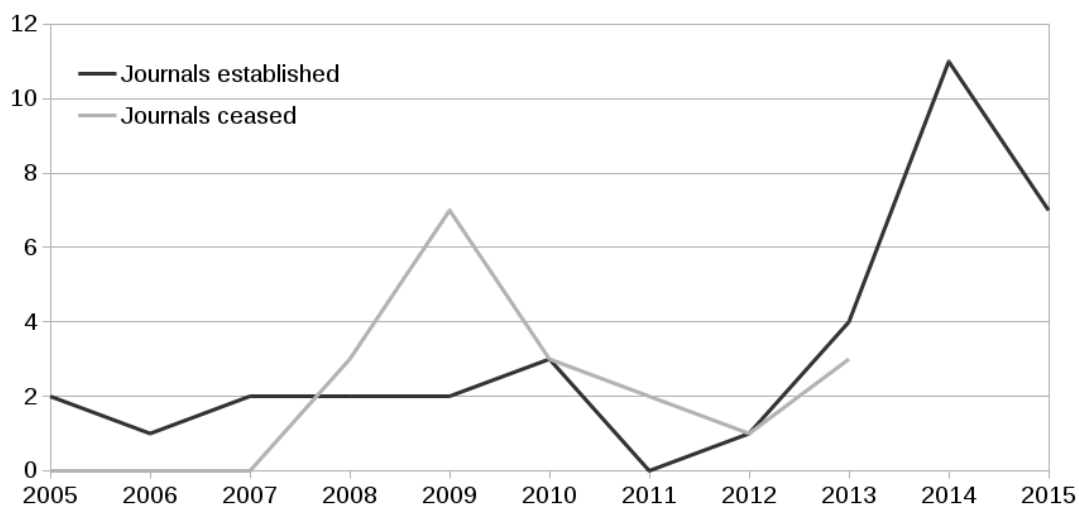
ONGOING WORK AND PRELIMINARY FINDINGS

Publication venues

In all 19 East-African countriesⁱⁱ combined, *Ulrichsweb*, known as standard journal database with high international coverage, lists 148 active scholarly journals in the selected field, with no journals in six countries (query on 2016-07-07).ⁱⁱⁱ 52 journal titles indicated a focus on applied research, and 37 started only after 2005 and therefore would have been too young to attract a wide audience of authors—all these have been excluded. With few exceptions, metadata on article level for these journals can only be found in *QIAPL*. Therefore, I decided to merge both lists. For this, I filtered the *QIAPL* journals according to my disciplinary and regional scope, looking only at the titles and the publisher (preferring research institutions), resulting in 103 journals. I found 27 titles matching the filtered *Ulrichsweb* list that have been filtered again because of inactivity in the relevant years 2008-2009 (12 journals). Many of the remaining titles in the *QIAPL* list are long ceased. Some are not included in *Ulrichsweb* (9), or are not labelled as "Academic / Scholarly" there (5), while still meeting my criteria. However, from a closer look at the article titles, I added them to my final journal list that now comprises of 26 titles, respectively 376 articles.

After these findings, for my specific field, *Ulrichsweb* can only be described as outdated database with insufficient coverage. I discovered that by far not all journals are actually active, although described as such: only seven have been publishing after 2013. It is interesting to see *when* they stopped publishing (Figure 2): the numbers reached a very clear peak in 2009. For the journal *Quest* that paused publishing in 2009 for almost three years, it is explained by a lack of funds during the financial crisis.^{iv} Compared to that, from my *Ulrichsweb* list filtered for basic research comprising 95 journals, with roughly half of them older than 25 years, very few new journals have been founded between 2005 and 2012. However, in 2014 not less than 11 journals started off.

Figure 2. Journal dying during financial crisis and recent journal growth. Ceased journals from final selection of 26 journals both in *Ulrichsweb* and *QIAPL*, compared to newly established journals according to *Ulrichsweb*, 95 journals filtered by discipline.



Indexing

None of the journals is listed in *Web of Science* or *Scopus*^v, but individual articles can be found, as references to other work, in *Google Scholar*, which therefore will provide the data for the planned citation analysis. At some point, ten of the 25 selected journals have been indexed—very selectively—in some bibliographic databases provided by *ProQuest* and *EBSCO*. In all but two cases this was discontinued, only one journal is still indexed, and the other has been indexed to the last issue that appeared. In roughly half of the cases, indexing most likely stopped because the journal paused. This discontinuity can almost always be explained with political events: the Ethiopian Civil War, or Daniel arap Moi's especially corrupt last presidential period. In other cases, indexing stopped although the journals appeared on time. Therefore, I plan to contact the journals as well as the database editors.

The few East-African SSH journals are badly covered by the main bibliographic databases that therefore manifest a hegemonic bias that leads to a corresponding bias of research conducted based on this foundation.

REFERENCES

- Archambault, É., & Larivière, V. (2010). The limits of bibliometrics for the analysis of the social sciences and humanities literature. In *World Social Science Report 2010* (pp. 251–254). Retrieved from http://crctcs.openum.ca/files/sites/60/2015/10/biblio_worldsocsciencereport.pdf
- Cloete, N., Maasen, P., & Baily, T. (Eds.). (2015). *Knowledge Production and Contradictory Functions in African Higher Education*. Somerset West: African Minds. Retrieved from <http://chet.org.za/download/file/fid/840>
- Esseh, S. K. S. (2011). *Strengthening scholarly publishing in Africa : assessing the potential of online systems*. Diss: University of British Columbia. Retrieved from <http://hdl.handle.net/2429/34184>
- Fari, S. A., & Ocholla, D. (2015). Comparative assessment of information and knowledge sharing among academics in selected universities in Nigeria and South Africa. *South African Journal of Libraries and Information Science*, 81(1). <http://doi.org/10.7553/81-1-1544>
- Fellessen, M., & Mählck, P. (2013). *Academics on the Move : Mobility and Institutional Change in the Swedish Development Support to Research Capacity Building in Mozambique*. Uppsala: Nordiska Afrikainstitutet. Retrieved from <http://urn.kb.se/resolve?urn=urn:nbn:se:nai:diva-1816>
- Mouton, J. (2010). The state of social science in sub-Saharan Africa. In *World Social Science Report 2010* (pp. 63–68). Retrieved from https://www.carnegie.org/media/filer_public/5c/91/5c91fb06-3ca7-47e8-9f78-bb2272357a1a/ccny_grantee_2010_social.pdf
- Nwagwu, W. (2005). Deficits in the visibility of African scientists: implications for developing information and communication technology (ICT) capacity. *World Review*

of Science, Technology and Sustainable Development, 2(3/4).

<http://doi.org/10.1504/WRSTSD.2005.007686>

Nwagwu, W. (2010). Cybernating the academe: Centralised scholarly ranking and visibility of scholars in the developing world. *Journal of Information Science*, 36(2), 228–241.

<http://doi.org/10.1177/0165551509358482>

Pouris, A., & Richter, L. (2000). Investigation into state-funded research journals in South Africa. *South African Journal of Science*, 96(3), 98–104.

Samuels, D. (2013). Book citations count. *PS: Political Science & Politics*, 46(04), 785–790.

<http://doi.org/10.1017/S1049096513001054>

Smart, P., & Murray, S. (2014). The status and future of the African Journal. In *The Future of the Academic Journal* (pp. 401–424). Chandos Publishing.

Tijssen, R. J. W. (2007). Africa's contribution to the worldwide research literature: New analytical perspectives, trends, and performance indicators. *Scientometrics*, 71(2), 303–327.

<http://doi.org/10.1007/s11192-007-1658-3>

-
- i I tested if the Kenyan Crisis (Dec 2007 to April 2008), which affected the whole region, had an impact on the publication quantity in a sample of the selected journals: compared to 2009, in 2008 even more articles have been published.
 - ii South Sudan is not listed in *Ulrichsweb*.
 - iii The final list of journals and all described steps can be followed up in the accompanying research data: East-African Social Sciences and Humanities Journals active in 2008-2009, <http://dx.doi.org/10.5281/zenodo.161541>.
 - iv Quest, <http://quest-journal.net>, last visited on 2016-07-13.
 - v With the exception of 21 articles from two journals.



Alphabetical co-authorship in the social sciences and humanities: evidence from a comprehensive local database¹

Raf Guns*

* raf.guns@uantwerpen.be

University of Antwerp, Centre for R&D Monitoring (ECOOM), Middelheimlaan 1, 2020 Antwerp, Belgium

ABSTRACT

We present an analysis of alphabetical co-authorship in the social sciences and humanities (SSH), based on data from the VABB-SHW, a comprehensive database of SSH research output in Flanders (2000-2013). Using an unbiased estimator of the share of intentional alphabetical co-authorship (IAC), we find that alphabetical co-authorship is more engrained in SSH than in science as a whole. Within the SSH, large differences exist between disciplines. The highest proportions of IAC are found for Literature, Economics & business, and History. Furthermore, alphabetical co-authorship varies with publication type: it occurs most often in books, is less common in articles in journals or in books, and is rare in proceedings papers. The use of alphabetical co-authorship appears to be slowly declining.

INTRODUCTION

A recurring debate related to research evaluation focuses on the question how multi-authored publications should be counted. For the sake of simplicity, we assume that one (single-authored) publication counts for one credit, but the argument is the same if the credit for a publication can vary based on other factors, such as citation count.

Two classic strategies for credit assignment are whole counting and fractionalized counting. Whole counting implies that each author receives one full credit for the publication, regardless of the number of authors. Fractionalized counting occurs when the credit is divided among the authors. In the simplest case, each of n authors receives $1/n$ credit.

More involved fractionalization strategies try to account for the relative contribution of each author, such that the author who has contributed the most to the paper receives the most credit, and so on. Most prior research starts from the assumption that the first author is the most important one, followed by the second one, and so on. Possible approaches include: only the first author receives credit (Cole & Cole, 1973; Persson, 2001), geometric counting (Egghe, Rousseau, & Van Hooydonk, 2000), and harmonic credit allocation (Hagen, 2008, 2013). Sometimes, other ways of determining author importance are used, such as recognition of the corresponding author or author contribution statements. Some research has pointed to the importance of the last author (Zuckerman, 1968) or the rise of 'equal first authors' (Hu, 2009).

Use of the byline order implies that authors deliberately order the byline according to relative contribution. A different criterion according to which authors can be ordered in the byline is

¹ I thank Tim Engels, Tryuken Ossenblok, Ronald Rousseau and Frederik Verleysen for their comments, which helped to improve the paper. This investigation has been made possible by financial support of the Flemish government to ECOOM.

simply by alphabetical order of their names. In that case, order of authors in the byline should not be used for credit allocation (Van Praag & Van Praag, 2008). Previous research (Frandsen & Nicolaisen, 2010; Levitt & Thelwall, 2013; Waltman, 2012) has already looked into the question to what extent alphabetical co-authorship occurs through time and in different disciplines. Major findings of these studies include:

- The number of authors per publication has increased over the last few decades;
- Some disciplines, like economics and mathematics, have a strong culture of alphabetical co-authorship whereas others do not;
- Overall, alphabetical authorship is declining.

These studies typically include (parts of) the social sciences and humanities or SSH in short. However, there are at least two reasons that suggest that, at this moment, no clear picture is available for the whole of SSH. First, most studies are concerned only with (parts of) the social sciences and much less with the humanities. Second, these studies are typically based on large international databases such as Web of Science (WoS). Since these databases only cover a limited subset of publications from SSH (Archambault et al., 2013), it is as of yet unknown to what extent alphabetical co-authorship is used in in SSH.

In this paper, we try to partially alleviate these limitations by using a comprehensive local database as data source. Specifically, we want to address the following questions:

1. To what extent do co-authored publications in SSH use alphabetical co-authorship?
2. How do different SSH disciplines differ in their use of alphabetical co-authorship?
3. How does alphabetical co-authorship vary with publication type?
4. How has alphabetical co-authorship evolved over the past decade?

DATA AND METHODS

Data

This paper uses data from the Flemish Academic Bibliographic Database for the Social Sciences and Humanities or VABB-SHW (www.ecoom.be/en/vabb). This is a comprehensive database of all peer reviewed publications by SSH researchers affiliated to a Flemish university from the year 2000 onwards (Engels, Ossenblok, & Spruyt, 2012; Ossenblok, Verleysen, & Engels, 2014).

The full data set covers the time period 2000–2013 and consists of bibliographic information for 59,560 peer reviewed publications with at least one author/editor affiliated to an SSH department at a Flemish university. There are 34,683 publications (journal articles, monographs, book chapters, or proceedings papers) with more than one author and 1,210 publications (edited books) with more than one editor. This subset of 35,893 publications (60.3% of all publications) is the basis of the present paper.

Table 1 provides additional details on the publication types and their counts in the data. Each publication is also assigned to one or more disciplines, according to the departmental affiliation of its authors; there are 9 disciplines in the humanities and 7 disciplines in the social sciences. In addition to those, the VABB-SHW has the general disciplines ‘Humanities general’ and ‘Social sciences general’.

Table 1. Overview of publication types in data set

	1 author/editor	>1 author/editor	Total
Journal articles	16,801	29,208	46,009
Monographs	621	287	908
Edited books	407	1,210	1,617
Book chapters	5,326	3,624	8950
Proceedings papers	512	1,564	2,076
Total	23,667	35,893	59,560

Publications in the VABB-SHW belong to two separate subcategories: VABB-WoS publications are publications that are also indexed in one of the WoS databases, whereas VABB-GP publications are not in WoS. The latter have been selected by the Authoritative Panel (Gezaghebbend Panel or GP), an independent body of academics charged by the Flemish Government with upholding the criteria for inclusion in the VABB-SHW, such as peer review. The majority of publications with more than one author/editor are VABB-WoS publications ($n = 20,298$ or 56.6%).

Methods

We can distinguish between intentional and accidental alphabetical co-authorship. Intentional alphabetical co-authorship (IAC) occurs when the authors have made a deliberate choice to put their names in alphabetical order on the publication, whereas accidental alphabetical co-authorship (AAC) occurs when the names are in alphabetical order but the byline order has been established using some other criterion (e.g., respective contribution to the end result).

First, we normalize names to account for slight variations in the way names are written. Spaces and other non-letter characters in author names are omitted, as well as accents. Full first names are reduced to the first letter. A name like ‘De Pré, Johan’, for instance, is normalized to ‘DEPRE, J’.

Next, we determine the number and proportion of publications (for the whole as well as per subgroup, e.g. per discipline or publication type) that have two or more authors and that are alphabetically authored. The resulting number and proportion count both IAC and AAC, and hence constitute an overestimation of the proportion of IAC. Other things being equal, for a publication with two authors that uses another criterion for the order in the byline, there is a 50% chance that the end result will be alphabetical and we end up with AAC. The probability of AAC rapidly decreases as the number of authors increases. In general, for a publication with n authors, the probability of AAC equals $1/n!$. Since we cannot directly measure IAC, we need to estimate it, accounting for the fact that, e.g., a paper with five authors in alphabetical order is much more likely due to IAC (99%) than a paper with only two (50%).

Waltman (2012) provides an unbiased estimator p for the proportion of IAC in a body of N publications, which is in turn based on a model by Van Praag & Van Praag (2008):

$$p = \frac{\sum_{i=1}^N \left(\frac{a_i - 1/n_i!}{1 - 1/n_i!} \right)}{N} \quad (1)$$

where n_i denotes the number of authors of publication i , and a_i is 1 if publication i is alphabetically authored and 0 otherwise. Note that p can only be applied to those publications that have more than one author. The estimator p typically ranges between 0 (no IAC) and 1

(complete IAC), although theoretically it can also yield negative values. A negative value would imply that authors deliberately seek out non-alphabetical orderings. Using formula (1), we can estimate the percentage of IAC in a body of co-authored publications, controlling for the number of authors on each individual publication.

RESULTS

Table 2 summarizes the results per discipline. All numbers reported here, as well as in the following tables, are based on the subset of publications with two or more authors or editors. Hence, the mean and median number of authors reported are higher than what has been reported in the literature, where single-authored publications are also counted (Ossenblok & Engels, 2015; Ossenblok et al., 2014).

We find large disciplinary differences in the proportion of alphabetical co-authorship. The proportion of alphabetical co-authorship ranges from 6.9% for Social health sciences to 48.1% for Literature. Although this is partially related to the number of co-authors per paper – the median for Social health sciences is 5, whereas for Literature this is only 2 –, the proportion of IAC for these two disciplines is still respectively lowest and highest. Remarkably, IAC is virtually non-existent in Social health sciences. The second and third places for highest proportion of IAC go to, respectively, Economics & business and History. Overall, we find that alphabetical co-authorship and IAC occur in almost every discipline, but in none of the disciplines are they the default or most used option. As can be seen from **Table 2**, alphabetical co-authorship and, especially, IAC occur more in the humanities than in the social sciences. While the humanities as a whole exhibit an IAC proportion of 12.0%, this is only 7.3% for the social sciences. Both percentages are, however, still significantly higher than the overall 3.7% reported by Waltman (2012), indicating that alphabetical co-authorship is more engrained in SSH than in science as a whole.

Table 2. Overview of results per discipline

	<i>N</i>	Mean authors	Median authors	Alphabetical order	IAC
Literature	563	2.89	2	48.1%	23.3%
Economics & business	7485	3.54	3	36.0%	18.7%
History	764	3.23	2	41.4%	16.3%
Criminology	1110	3.31	2	39.0%	14.9%
Philosophy	1482	3.89	3	34.7%	13.7%
Linguistics	2341	3.41	3	35.5%	12.5%
Theology	401	3.72	3	33.2%	11.8%
Humanities general	1434	3.43	3	33.6%	11.2%
Law	2508	2.82	2	42.2%	10.9%
Political sciences	1463	2.89	2	38.8%	10.4%
Communication studies	1168	3.71	3	32.1%	9.9%
Sociology	2228	4.32	3	24.5%	6.8%
History of arts	871	4.01	4	24.2%	6.5%
Archaeology	599	5.03	4	17.4%	6.2%
Social sciences general	5794	4.89	4	16.7%	3.8%
Educational sciences	2327	3.84	3	20.6%	3.2%
Psychology	5638	4.71	4	15.7%	1.9%
Social health sciences	8023	6.65	5	6.9%	0.4%

Next, we turn to the question to what extent alphabetical co-authorship varies with publication type. The results are summarized in **Table 3**. Monographs and edited books are quite different from the other three publication types, in that they display a much larger share of IAC. The high percentage of IAC for edited books confirms previous research (Ossenblok & Engels, 2015; Ossenblok, Guns, & Thelwall, 2015) that edited books and editors are quite different from other publication types and their authors. The low proportion of IAC for proceedings papers is somewhat unexpected and deserves further investigation. Journal articles and book chapters exhibit a comparable proportion of alphabetical co-authorship and IAC.

Table 3. Overview of results per publication type

	<i>N</i>	Mean authors	Median authors	Alphabetical order	IAC
Journal articles	29,195	4.53	3	24.2%	7.6%
Monographs	284	3.25	3	43.3%	21.4%
Edited books	1,209	3.08	3	49.0%	30.0%
Book chapters	3,623	3.59	3	29.5%	9.2%
Proceedings papers	1,564	4.66	4	15.0%	2.6%

Given the large amount of journal articles, the results in **Table 3** cannot show all variation that exists within this publication type. We therefore look into the question to what extent alphabetical co-authorship varies with whether or not a journal is indexed in WoS. As was mentioned earlier, the majority of co-authored publications are VABB-WoS publications. Indeed, we know from prior research that more internationally oriented literature – often indexed in databases like WoS or Scopus, authored in English, etc. – tends to have more

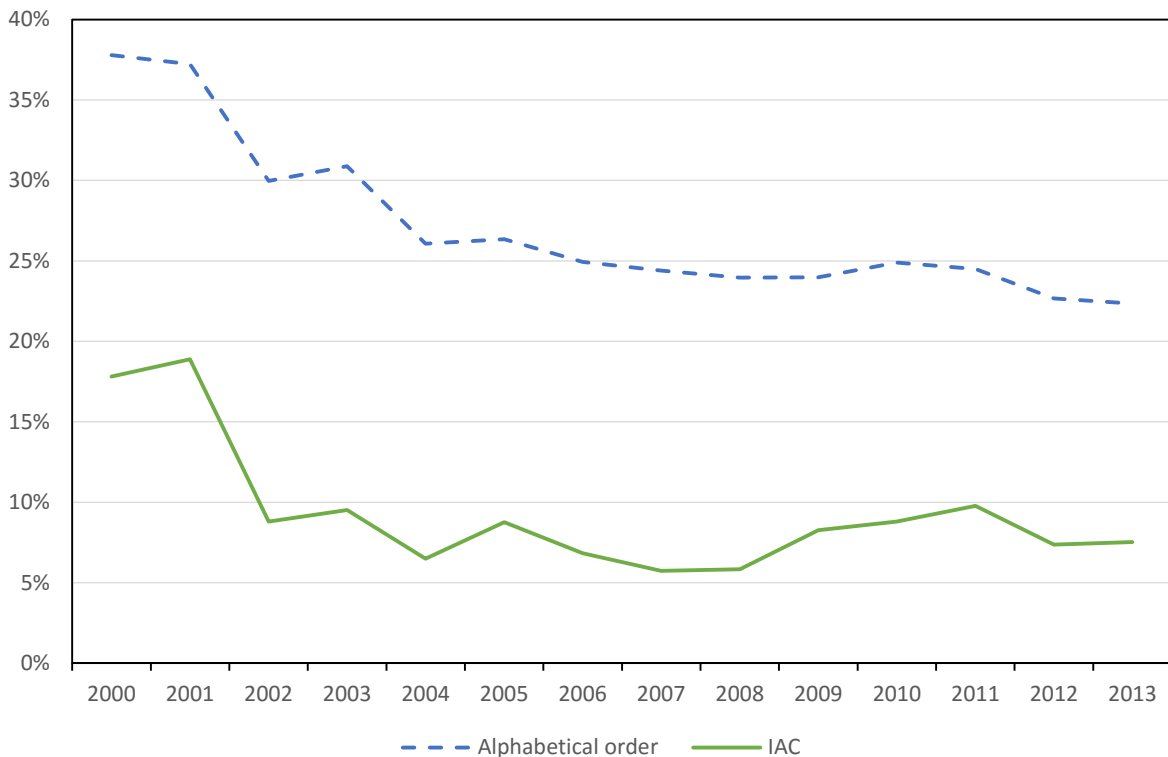
authors (Ossenblok et al., 2014). This can also be seen from the mean and median number of authors for VABB-WoS compared to VABB-GP journal articles (**Table 4**). The difference between WoS-indexed and non-WoS-indexed publications is even more striking when we consider byline order. Both alphabetical co-authorship and IAC occur about twice as frequently in publications that are not indexed in WoS. This discrepancy can be seen in both the humanities and social sciences, although the difference is most pronounced for the social sciences. This strongly suggests that alphabetical co-authorship within SSH is most prevalent in locally or regionally oriented literature, even if we account for the fact that the number of co-authors is typically lower there.

Table 4. Comparison of VABB-WoS and VABB-GP journal articles

		<i>N</i>	Mean authors	Median authors	Alphabetical order	IAC
VABB-GP	Humanities	3,996	2.83	2	42.0%	11.8%
	Social sciences	6,732	3.47	3	32.4%	10.8%
	All	9,834	3.22	3	35.8%	10.8%
VABB-WoS	Humanities	2,759	4.31	3	26.2%	7.2%
	Social sciences	17,423	5.35	4	17.0%	5.6%
	All	19,361	5.19	4	18.4%	5.9%

Figure 1 displays the evolution of alphabetical co-authorship and IAC between 2000 and 2013. After a relatively strong decline over the first five years, the proportions remain fairly stable around 7 to 8%. This is broadly in line with Waltman’s (2012) finding that alphabetical co-authorship is declining.

Figure 1. Evolution of proportions of alphabetical order and IAC (2000–2013)



DISCUSSION AND CONCLUSION

We have shown that alphabetical co-authorship occurs more in SSH than in other fields of science and that authors in the humanities choose an alphabetical ordering of their names in the byline more frequently than in the social sciences. At the same time, these overall findings should not obscure the fact that there exists quite a bit of variety within the social sciences and humanities. The highest proportions of IAC are found for Literature, Economics & business, and History. The use of alphabetical co-authorship is declining over time.

The proportions of alphabetical co-authorship and IAC in Economics & business we find are clearly lower than those reported in previous research (Frandsen & Nicolaisen, 2010; Levitt & Thelwall, 2013; Waltman, 2012), where, for instance, the proportion of alphabetical co-authorship was roughly around 75%. The main explanation lies in the way publications are assigned to disciplines in the VABB-SHW. A publication is assigned to a discipline if at least one of its authors belongs to an administrative unit that is assigned to this discipline. In the case of Economics & business this has led to a situation where many publications are only tangentially related to the core fields of Economics & business. This was verified by looking at the subset of WoS publications from this discipline. If we restrict the WoS publications from Economics & business to those with WoS SC Economics (20.3%), we find an IAC proportion of 48.4%. However, this is still lower than the 58% reported by Waltman (2012).

As for publication types, alphabetical co-authorship occurs most often in books, is less common in articles in journals or in books, and is rare in proceedings papers. Articles in WoS journals exhibit less alphabetical co-authorship than those in GP-selected journals. The differences between publication types cannot be explained solely through disciplinary preferences for certain publication types (Nederhof, 2006), given the high percentage of IAC for edited books and monographs.

Ideally, the order of authors in the byline should not be used for credit assignment in those disciplines where the proportion of IAC exceeds a certain threshold. Even if the proportion of IAC stays well below the threshold, this does not necessarily imply that the byline order accurately reflects each author's contribution.

REFERENCES

- Archambault, É., Vignola-Gagné, É., Côté, G., Larivière, V., & Gingras, Y. (2013). Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. *Scientometrics*, 68(3), 329–342.
- Cole, J. R., & Cole, S. (1973). *Social stratification in science*. Chicago: U.Chicago Press.
- Egghe, L., Rousseau, R., & Van Hooydonk, G. (2000). Methods for accrediting publications to authors or countries: Consequences for evaluation studies. *Journal of the American Society for Information Science*, 51(2), 145–157.
- Engels, T. C. E., Ossenblok, T. L. B., & Spruyt, E. H. J. (2012). Changing publication patterns in the Social Sciences and Humanities, 2000–2009. *Scientometrics*, 93(2), 373–390.

- Frandsen, T. F., & Nicolaisen, J. (2010). What is in a name? Credit assignment practices in different disciplines. *Journal of Informetrics*, 4(4), 608–617.
- Hagen, N. T. (2008). Harmonic allocation of authorship credit: source-level correction of bibliometric bias assures accurate publication and citation analysis. *PloS One*, 3(12), e4021.
- Hagen, N. T. (2013). Harmonic coauthor credit: A parsimonious quantification of the byline hierarchy. *Journal of Informetrics*, 7(4), 784–791.
- Hu, X. (2009). Loads of special authorship functions: Linear growth in the percentage of “equal first authors” and corresponding authors. *Journal of the American Society for Information Science and Technology*, 60(11), 2378–2381.
- Levitt, J. M., & Thelwall, M. (2013). Alphabetization and the skewing of first authorship towards last names early in the alphabet. *Journal of Informetrics*, 7(3), 575–582.
- Nederhof, A. J. (2006). Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A Review. *Scientometrics*, 66(1), 81–100.
- Ossenblok, T. L. B., & Engels, T. C. E. (2015). Edited books in the Social Sciences and Humanities: Characteristics and collaboration analysis. *Scientometrics*, 104(1), 219–237.
- Ossenblok, T. L. B., Guns, R., & Thelwall, M. (2015). Book editors in the social sciences and humanities: an analysis of publication and collaboration patterns of established researchers in Flanders. *Learned Publishing*, 28(4), 261–273.
- Ossenblok, T. L. B., Verleysen, F. T., & Engels, T. C. E. (2014). Coauthorship of journal articles and book chapters in the social sciences and humanities (2000-2010). *Journal of the Association for Information Science and Technology*, 65(5), 882–897.
- Persson, O. (2001). All author citations versus first author citations. *Scientometrics*, 50(2), 339–344.
- Van Praag, C. M., & Van Praag, B. M. S. (2008). The Benefits of Being Economics Professor A (rather than Z). *Economica*, 75(300), 782–796.
- Waltman, L. (2012). An empirical analysis of the use of alphabetical authorship in scientific publishing. *Journal of Informetrics*, 6(4), 700–711.
- Zuckerman, H. A. (1968). Patterns of Name Ordering Among Authors of Scientific Papers: A Study of Social Symbolism and Its Ambiguity. *American Journal of Sociology*, 74(3), 276–291.



Aligning research assessment in the Humanities to the national Standard Evaluation Protocol Challenges and developments in the Dutch research landscape

1st organizer: Ad Prins, Support in Research Management, Pausgang 4, 9712 EZ Groningen info@adprins.nl

2nd organizer: Jack Spaapen, Royal Netherlands Academy of Arts and Sciences, PO Box 19121, 1000 GC Amsterdam, jack.spaapen@knav.nl

3rd organizer: Frank van Vree, Dean of the Humanities Faculty of the University of Amsterdam, Bungehuis, Spuistraat 210, 1012 VT Amsterdam, F.P.I.M.van.Vree@uva.nl

(Members of the Steering group for *Quality and Impact Indicators for the Humanities*, a project initiated by the Deans of Faculties of Humanities in the Netherlands)

ABSTRACT

Purpose and intended audience

The purpose of this session is a debate about innovation in comprehensive methods for the assessment of humanities research. Input will come from preliminary outcomes of an ongoing project in the Netherlands to find adequate indicators for humanities research that will fit in the national Standard Evaluation Protocol. The project includes processes of 'bottom up' data collection (that is, with input coming from the research community) and discussion with Humanities researchers, investigating the specific characteristics of publication and communication cultures in the Humanities, and the prospects for the use of quantitative and qualitative indicators.

The expected outcome is threefold, first we hope that the materials to be presented will enable a comparison with similar initiatives in the UK, Flanders and Norway. Second, we foresee that the session will strengthen the final outcomes of the Dutch project in view of the international feedback, and third, we hope to further the discussion about comprehensive assessment and the use of indicators in the Humanities.

Intended audiences are both users and producers of humanities research, that is researchers and stakeholders in both scientific communities and societal contexts.

Proposed activities

The session will entail three sections (1) presentation and discussion of the Dutch context of research evaluation, set by the national evaluation protocol (SEP 2015-2021), with a focus on possibilities and constraints for research evaluation in the Humanities. (2) Discussion about preliminary outcomes of the project, including information about publication cultures, discussion with the research community about pros and cons of particular indicators for quality and impact. (3) A discussion about the common ground in the different approaches in the Netherlands, Norway, Flanders and the UK, and possibly other countries.

We aim at organizing the session as a mini living lab, that is as an event where users and designers of evaluation processes co-create an outcome. We will invite attendees from different countries to prepare short comments on questions and issues that we will distribute in advance. Each of the three sub-sections will be 25 minutes, introductions to the sections will be 5-10 minutes, discussion time 15- 20 minutes per section. At the end there will be 10-

15 minutes times to go over the results of the whole session. We will prepare a report that will serve as input for the Dutch project, and will be distributed to attendees of the session for further comments. Further dissemination to those interested is also intended.

Relevance to the conference / significance to the field

Science and technology indicators in use today are predominantly based on publication and communication patterns in the STEM fields. Therefore, they are often not adequate for SSH fields, because the publication and communication patterns are different. This session explores new ways to work with indicators that are better representing communication and publication patterns in the Humanities, including new insights of bibliometric characteristics based on Google Scholar data.

Novelty

Our approach is user-oriented, that is, it is a bottom-up approach, it includes stakeholders (researchers and users) in designing new methods for quality and impact assessment in humanities research.

Length: 90 minutes

Preferred number of participants: 35

Requests: a projector

References

Dávidházi (ed), *New Publication Cultures in the Humanities*, AUP 2014

KNAW, NWO, VSNU, (July 2015), Standard Evaluation Protocol 2015-2021, protocol for research assessment in the Netherlands, The Hague: KNAW, NWO, VSNU

Royal Netherlands Academy of Arts and Sciences (2011), *Quality Indicators for Research in the Humanities*, Amsterdam: KNAW

Giménez-Toledo, E., J. Mañana-Rodríguez, T. Engels, P. Ingwersen, J. Pölönen, G. Sivertsen, F. Verleysen and A. Zuccala, *The Evaluation of Scholarly Books as a Research Output. Current Developments in Europe*, undated

Spaapen, Jack, and A.A.M. Prins, *From research impact assessment to contextual evaluation, Philosophy and Technology*, forthcoming,

PROPOSAL

In the Netherlands, since 2003 most publicly funded research is evaluated every six years according to the Standard Evaluation Protocol (SEP). After each six year cycle, the protocol is reviewed and updated where necessary. The SEP runs under the auspices of three organisations: the Association of Universities in the Netherlands (VSNU), the Netherlands Organisation for Scientific Research (NWO), and the Royal Netherlands Academy of Arts and Sciences (KNAW). The current SEP was introduced in 2015 and will run up to 2021.

The Standard Evaluation Protocol (SEP) describes the methods used to assess research conducted in the Dutch universities and the research institutes of NWO and KNAW. Also, a number of independent publicly funded research institutes outside these three organisations, use the SEP to evaluate their work.

In the SEP 2015-2012 judgement is based on three assessment criteria: research quality, relevance to society, and “viability” (the extent to which the unit is equipped for the future). Assessment committees are international, and have to deliver a verdict both in text (qualitative) and in categories (quantitative). The four possible categories are “excellent”, “very good”, “good” and “unsatisfactory”. Two further aspects are to be considered: PhD programmes (including those at the national research schools) and research integrity. Here, the committee limits itself to a qualitative assessment.

Indications how to assess research quality are given in the SEP as follows: The committee assesses the quality of the unit’s research and the contribution that research makes to the body of scientific knowledge. The committee also assesses the scale of the unit’s research results (scientific publications, instruments and infrastructure developed by the unit, and other contributions to science).

regarding the relevance to society, the committee is asked to assess the quality, scale and relevance of contributions targeting specific economic, social or cultural target groups, of advisory reports for policy, of contributions to public debates, and so on. The point is to assess contributions in areas that the research unit has itself designated as target areas.

The research units to be evaluated are asked to deliver evidence of their performance in the last 6 years for the following six evaluation categories, divided over two main evaluation criteria: scientific quality and societal relevance, see schedule:

	Scientific quality	Relevance to society
Demonstrable output	Sc. articles (refereed vs. non-refereed) Sc. Books, classification of publishers Other research outputs (instruments, infrastructure, datasets, softwaretools, designs) Dissertations	(policy) reports Articles in professional journals Other output (instruments, infrastructure, datasets, softwaretools, designs) Outreach-activities, public lectures, exhibitions,
Demonstrable use	Citations Use of datasets, softwaretools, etc. by peers Use of research facilities by peers Reviews in scholarly journals	Patents/licences Use of research facilities by societal partners Projects with societal partners Contract research
Demonstrable recognition	Scientific prizes Personal sc. subsidies Invited lectures Membership of sc. committees, editorial boards, etc.	Public prizes Valorisation funding Positions paid for by public parties Memberships of public advisory bodies

The indicators mentioned in the table above are meant as examples. The basic idea of SEP is that it presents an overall framework, the six cells in the table, but that the indicators within the table are to be decided upon by the different fields of science. The social sciences might consider other indicators than for example the engineering fields.

For the humanities, the deans of all faculties in the Netherlands decided to start a project that aimed at developing indicators for all six categories that would be adequate for all or most subfields in the area of humanities. The project is conducted in the years 2015 and 2016. Our proposal is to present and discuss this project at the STI ENID conference.

THE DEVELOPMENT OF INDICATORS FOR HUMANITIES RESEARCH IN THE NETHERLANDS

The main goal of the project is to develop an assessment system that adequately represents scientific work done in the humanities and fits in the SEP framework. It is oriented towards both qualitative and quantitative indicators, and it considers the development of such systems in other European countries, in particular in Norway, Flanders and the UK. The project is conducted by a core group established by the deans, but as a rule it involves both the humanities faculties in the Netherlands and the research schools. The latter are national units in which research MA students are supervised, and PhD's are trained. Some of these national research schools have a broad orientation, and exist for a long time, others have been installed more recently. They operate within the university system in a semi-independent mode.

There are a number of challenges in this project. Firstly, since the humanities are a field with lots of different fields, ranging from the more classical disciplines of languages, history, philosophy, and theology to the more recent areas in culture and media studies, serious and fun gaming and since a few years also digital humanities, it will be difficult to come up with something that satisfies all these fields. Secondly, many fields in the humanities do not have publication traditions comparable with the so-called STEM fields, that are oriented towards a limited set of international journals. In many fields, publications of books, or book chapters, are more common. Thirdly, and connected to the previous one, the development of robust databases has not been as strong as in the case of STEM fields. An analysis of the ISI database or of google scholar supports this point. of the project is And finally, the kind of output of humanities research is overall less easy to catch in concrete terms, it is oriented more towards increasing knowledge, raising insight and awareness on certain topics or issues, than it is towards solving concrete questions.

A central question then is whether it is at all possible to develop a system that compares even to a certain extent to what has been developed for the STEM fields, or that the humanities want a completely different approach to evaluation.

Having said that, it is clear that humanities need a system that is as robust as possible, if only because the various funding systems work out negatively for those fields that are not able to deliver substantial evidence for their quality and relevance. And this is even more pressing since in various countries, like Japan or the USA, governments openly doubt the value of humanities research, and suggest dire consequences for the funding of these fields.

In all simplicity, the project entails two phases, one in which we review the publication (and other output) cultures in various humanities fields. In this we identify, together with the research communities, the most common communication channels for research output, whether this is through journal articles, books, or other forms of output. The second phase will be an analysis based on the results of the first phase of the possible quantitative and qualitative indicators. This analysis also entails an assay of a number of databases that might be relevant for the output of humanities research (such as google scholar, or specialised databases).

The projected outcome of the project is a model fit for the humanities that would

- value on a par scientific quality and societal value of research
- be multifunctional: it would serve not only to assess individual performance, but also the performance of groups and programs
- be useful to discuss research policy, including the improvement of societal relevance of research and career policy for young researchers;
- include suggestions for the role of altmetrics and of open access publications (P. Dávidházi (ed), *New Publication Cultures in the Humanities*, AUP 2014);

CHAPTER 12

Society, Participation and Culture



Operationalizing RRI: Relational Quality Assessment & Management Model for Research and Innovation Networks (REQUANET)

Julieta Barrenechea^{*}, Andoni Ibarra^{**}

^{*}julieta.barrenechea@euskampus.eu

Miguel Sanchez-Mazas Chair UPV/EHU, Av. Tolosa 70, Donostia-San Sebastián, 20018 (Spain)

^{**}andoni.ibarra@ehu.eus

Miguel Sanchez-Mazas Chair UPV/EHU, Av. Tolosa 70, Donostia-San Sebastián, 20018 (Spain)

INTRODUCTION

The current European framework for research and innovation, *Horizon 2020*, is articulated through the Responsible Research and Innovation (RRI) approach, promoting new dynamics of science and society (Owen et al., 2012) that require guidance of scientific production and public policy with a focus on developing multi-agent, transdisciplinary, mission-oriented solutions for local and global challenges.

Promoting research and innovation through *science with and for society* dynamics questions traditional views on scientific quality, associated almost exclusively with criteria of academic excellence. Rather, assessing through RRI perspective should help to make scientific activity permeable to a broad range of considerations and, accordingly, research and innovation dynamics would be more contingent or debatable.

From this viewpoint, we consider that assessing and managing RRI-driven science constitute a peripheral cognitive space. Two contributions can be highlighted regarding this task: the *Report from the Expert Group on Policy Indicators for RRI* (EC 2015) and Wickson/Carew (2014). Both contributions define quality criteria and indicators for promoting and monitoring RRI processes, but they limit the intrinsic openness and contingency of innovation dynamics by establishing RRI in a substantive way through a number of “quality criteria” or “key areas”.

Understanding responsibility of science and innovation in more relational or systemic terms needs to consider responsibility throughout the whole process of innovation, including the values, motivations and expected benefits behind innovations. Consequently, RRI research and innovation network’s monitoring and assessment should make contingent the assumptions, values and dimensions underlying innovation systems, and be used to appraise innovations in terms of their processes and/or expected benefits or “right impacts”.

This contribution proposes operationalizing RRI through a management and assessment model of networks that takes into account the contingent character of RRI. The *Relational Quality Assessment & Management Model for Research and Innovation Networks* (REQUANET) understands RRI from a network perspective, consisting of stakeholders working together during “the whole process of innovation” (von Schomberg 2015). This model focuses on interactions with a relational (rather than substantial) view of research and innovation networks.

Examples will be shown of its experimental application involving the Cooperative Research Center in Biosciences –CRC bioGune- in the Basque Country.

REQUANET MODEL FRAMEWORK

REQUANET combines two concepts, (i) *relevant connectivity* and (ii) *socio-technical robustness*, into a third concept, (iii) *relational quality*. The link between these three concepts combines relational and dynamic elements with a projective focus that takes account of the two categorical RRI components: the requirements of *co-responsibility* and a prospective approach based on *social desirability*.

(i) *Relevant connectivity* refers to the *science in and for society* constraint proposed under the RRI approach, where science and society are still perceived as differentiated spheres. It identifies scientific cooperation networks as it searches for the social relevance of research, i.e. for the integration of social concerns, perspectives, challenges and priorities.

(ii) *Socio-technical robustness* expresses the states of socio-technical configurations of such networks, as a driver promoting *science in and for society* dynamics.

Socio-technical robustness is the goal of scientific cooperation networks. It considers the three conditions identified in RRI: (i) inclusion of a variety of perspectives, (ii) mutual responsibility in prospective and *social desirability* terms and (iii) a learning dimension marked by the *mobilization of reflexive capital*.

(iii) *Relational quality* is associated with states of socio-technical integration that are products of the learning capacity of science cooperation networks in their processes of socio-technical robustness deployment and thus of responsible research and innovation.

REQUANET LEVELS AND METHODS

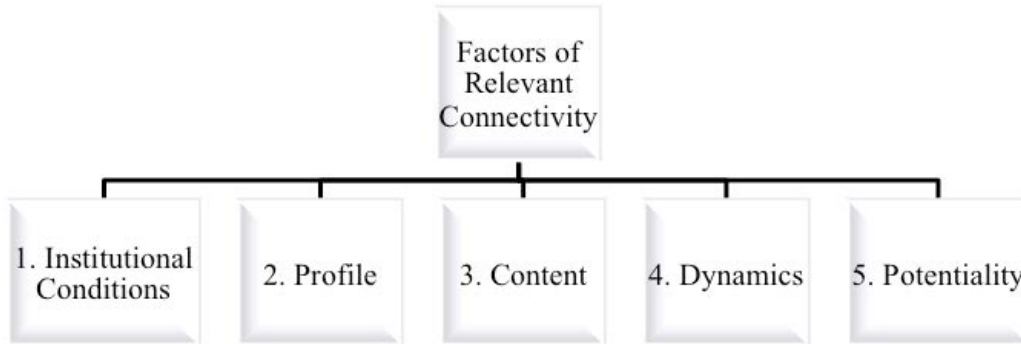
REQUANET incorporates methods in 4 levels that help in promoting or accompanying the multiplication of pertinent and horizontal processes of science-society interaction. The model proposes processes oriented towards overcoming the *science in and for/towards society* dynamic for the sake of greater socio-technical integration through the deployment of (socio-technical) robustness.

REQUANET is based on a Conceptual Framework and four levels:

- *Level 1: Descriptive evaluation of conditions for relevant connectivity*
- *Level 2: Analysis of associations between factors of relevant connectivity*
- *Level 3: Calculation of relevant connectivity index and subindexes*
- *Level 4: Deployment of socio-technical robustness*

The Conceptual Framework defines five relevant connectivity factors.

Figure 1. Factors of Relevant Connectivity



The factors are operationalized in indicators to identify the strengths and weaknesses of the units under analysis (a project, a Research Center, ... as a network). As an example, we present *Profile* (Figure 2) and *Potentiality* (Figure 3) subfactors.

Figure 2. Profile subfactors

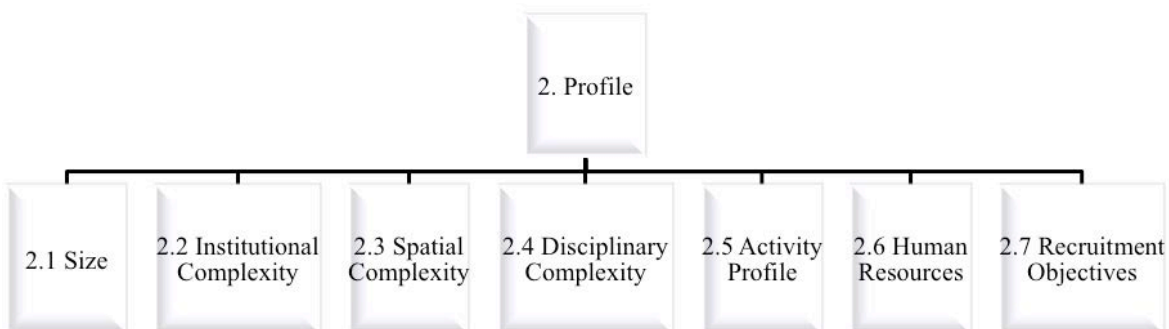
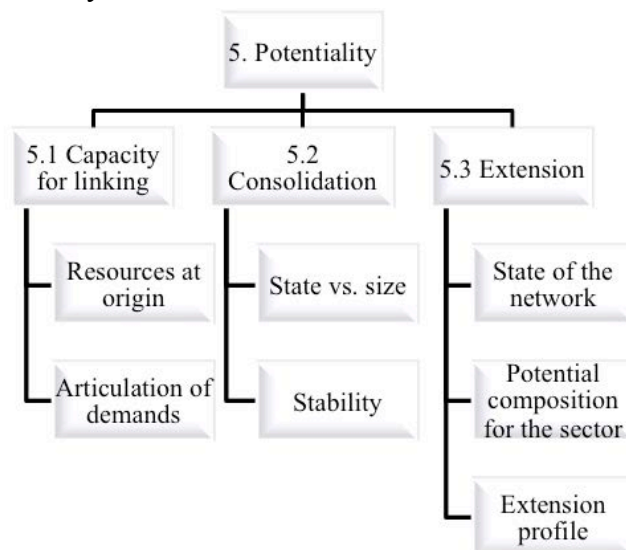


Figure 3. Potentiality subfactors and indicators



Indicators of the *Extension* subfactor (*Potentiality* factor) are for instance:

State of the network: ratio between links already formed and the number of identified potential links.

Potential composition for the sector: based on the diversity of the sectors of links.

Extension profile: based on the diversity of the extension profiles of the potential links.

REQUANET combines a quantitative approach present in the first three levels with a qualitative approach included as a specific moment in Level 3, and which completely defines Level 4. In the quantitative approach, the main structuring criterion is a direct relation between disciplinary, sectorial, geographical, etc. *diversity* and *connectivity*: the greater the diversity, the higher the level of relevant connectivity, in short, the greater the level of socio-technical integration; the decisive factor in relational quality. A distinguishing feature of the qualitative approach is that, besides being analytical, it is dynamic, as it concentrates on the inclusive mobilization of actors in the network and their reflexive capital.

Analytical tools for identifying the starting conditions for relevant connectivity and its potential outreach are combined with tools suitable for including and empowering the actors in the potential network. As well as integrating qualitative and quantitative focuses, REQUANET combines four approaches: contextual, descriptive, interactive and projective.

1. Institutional Conditions	• Contextual approach
2. Profile	• Descriptive approach
3. Contents	• Descriptive approach
4. Dynamic	• Interactive approach
5. Potentiality	• Projective approach

CRC BIOGUNE, THE CONTEXT OF REQUANET EXPERIMENTAL APPLICATION

CRC bioGUNE is located in the Bizkaia Technology Park, Basque Country and began operating in 2004 within the framework of the bioBask 2010 bioscience development strategy. The organizational model draws on the *CRC Program* developed in the 1990s by the Australian Government.

Its mission is to provide the system with new scientific capacities and to bring together and coordinate the activities of the actors intervening in a single strategic area. It is structured in

two levels: a) *physical* CRC, responsible for scientific research in the strategic areas of bioscience, and b) *virtual* CRC, a network of actors joining together through effective articulation in research projects.

The CRC model provides a solid organizational context for our study, primarily because:

- a) connectivity forms an explicit part of its mission,
- b) it seeks to articulate STI public policy, regional socio-economic relevance and research agendas,
- c) it aims to contribute to the development of the science-society collaborative culture and
- d) the development cycle of its projects provides for moments when an extended community of social players can be included.

On the basis of the RRI approach, critical analysis of this organizational model's *relational scope* can be conducted according to the four following points:

- a) it adopts a meso-level of inclusion of social desirability, as it integrates STI public policy guidelines into its research agendas and regional socio-economic priorities concerning biosciences, which have been defined by processes of strategic reflection with multisectorial representation,
- b) it promotes prompt science-industry interactions primarily oriented towards the assessment of knowledge *towards* the market. Without detracting from the efficiency of the model for this objective, from the RRI perspective, CRC represents a limited science-society dynamic; strictly speaking, it would fail to reach a level of "*science for and with society*",
- c) it is paradigmatic in terms of the effective development of collaborative culture and competences,
- d) it may boost the science-society dynamic as it includes actors when problematizing and defining project agendas.

In short, we see potential for the CRC organizational model to take its dynamic towards more far-reaching relational frontiers, primarily based on two drivers: a) *broadening inclusion* on the basis of greater diversification of the community at key points in the cycle, and b) *developing higher levels of reflexivity and network learning*, the explicit mission to develop connectivity involves network self-representation facilitating this challenge.

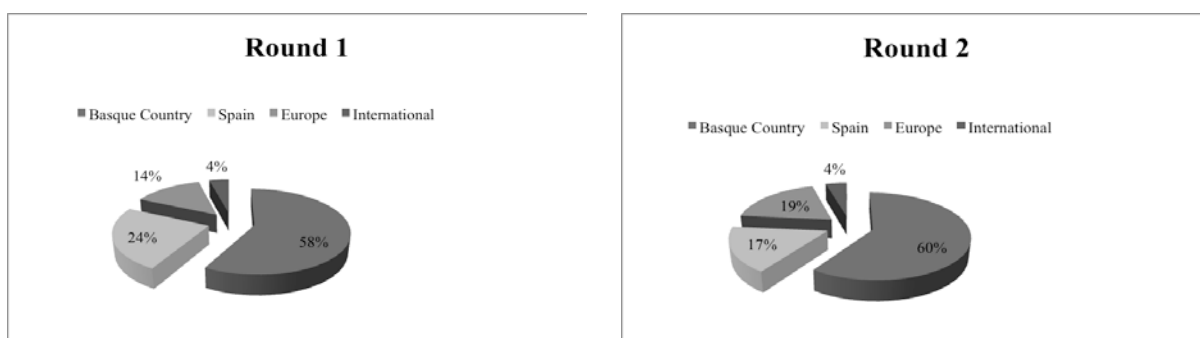
REQUANET APPLICATION LEVELS AND EXEMPLIFICATION OF RESULTS

The conceptual framework for relevant connectivity has been operationalized using a semi-structured questionnaire, applied to the coordinators of five collaborative projects underway at CRC bioGUNE in two consecutive rounds spacing one year. The REQUANET levels are described below and some results exemplified.

Level 1: *evaluation of relevant connectivity conditions*. Descriptive statistics are applied to unearth conditions for relevant connectivity in scientific and innovative networks for each factor and sub-factor making up the Conceptual Framework for Relevant Connectivity.

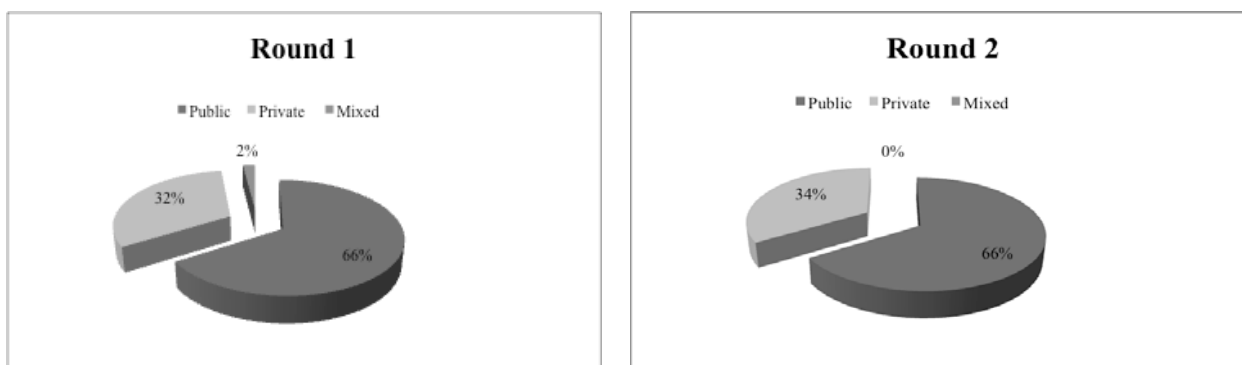
For example, the compared results of the two rounds regarding two Profile dimension indicators are cited.

2.2 Institutional Complexity - 2.2. b Links according to sector



The *institutional profile* of the network shaping the five projects shows that 66% of the links are established with public institutions, 32% with private and just 2% with mixed institutions. There have been no modifications to the network's institutional profile between rounds.

2.3 Special complexity - 2.3.a Distribution of links according to area of origin



The distribution of links according to area shows that the percentage of links with the Basque Country has increased slightly (58% R1 and 60% R2), and a more significant decrease can be seen in links with Spain (24% R1 and 17% R2), and a 5% increase in links with Europe (14% R1 and 19% R2).

Level 2: *analysis of associations between factors of relevant connectivity*. The idea is to cross the exploratory variables to establish inter- and intra-factor links focusing on *Profile* and *Potentiality*. It applies correlation quotients in a management dynamic based on the exploration of links and impacts between variables.

A more significant body of data would be required to apply this level. It is therefore a proposal yet to be exploited. Topics are defined from which certain associations are proposed between variables and the hypothetical relation sought to be explored is highlighted.

For example, Topic a): Profile impact (the main expression of connectivity) on the potentiality of the socio-technical network in order to develop relevant connectivity.

Topic a.1) Exploratory hypotheses referring to: “impact of size on potential projection”

“impact of size on the capacity for linking”

2.1 a) Size – (6.1 b) The link’s resources at origin

2.1 a) Size – (6.1 c) Articulation of demands according to sector

“impact of size on the capacity for extension”

2.1 a) Size - (6.3 a) Number of potential nodes

Topic a.2) Hypothesis “the greater the institutional complexity, the greater the potential network connectivity”

“Greater IC - greater capacity for linking”

2.2 Institutional complexity – 6.1 a) The link’s resources at origin

2.2 Institutional complexity – 6.1 b) Articulation of demands according to sector

“Greater IC, greater capacity for extension”// “greater social pertinence”

2.2 Institutional complexity – 6.3 b) Potential composition according to sector (diversity and/or identifying if a particular sector stands out in accordance with predominant 2.2 sectors)

2.2 Institutional complexity – 6.3 d) Extension profile

Level 3: combines the quantitative approach of the *calculation of the relevant connectivity index (RCI) and subindexes (subRCI)* with the inclusion of the first qualitative and prospective moment of inclusive reflexivity.

Comparison of the results from the index and subindexes between rounds shows the type of connectivity level monitoring that can be conducted with this model. Using the RCI results between rounds, evolution of the global relevant connectivity level is analysed, which is a single summarizing measurement for the scientific collaboration network as a whole. Using the partial results, subRCI, the evolution of relevant connectivity for each connectivity factor can be discriminated.

The qualitative moment is defined by the possibility of including an extended community from the scientific and innovative network to perform a weighting exercise concerning factors

relating to strategic objectives and priorities to improve network connectivity. Thus, REQUANET takes into account two requirements that allow the *science for and with society* level defined by RRI to be achieved: promoting (i) processes of strategic inclusive reflexivity, and (ii) inclusive empowerment.

How the following are obtained from the factor weighting exercise:

Result	Utility	Presentation mode
Individual weighting of factors for each participant (which may represent an institution or collective).	Identifying each institution or collective's priorities regarding connectivity factors and their specific needs or goals. Monitoring.	Table showing percentage data
Collective weighting of factors (average of weightings)	Identifying collective priorities in relation to relevant connectivity goals for the scientific collaboration network. In combination with the index results, actions oriented towards achieving these goals can be defined and planned. Monitoring.	Table showing percentage data Pie chart

Having weighted the factors, the relevant connectivity index can be calculated. The following results were obtained from the case studied:

Results	Relevant Connectivity Subindexes according to Factor (subRCI)					Relevant Connectivity Index (RCI)
	Institutional Conditions	Profile	Content	Dynamic	Potentiality	
Round 1	15%	53.97%	48%	58.86%	62.96%	45.97%
Round 2	20%	54.30%	68%	45.05%	73.85%	47.79%
Deviation R1 Standard	1.37	0.68	0.91	1.58	0.41	0.99
Deviation R2 Standard	1.12	0.71	1.20	0.54	0.94	0.90

Analysis of standard deviation enables us to ascertain if there are extreme values dispersing the average results. Taking into account that the RCI is based on the sum of weighted averages, this analysis is important so as to elucidate the representation of the RCI and subRCI results.

For the global RCI level, standard deviation helps to analyse the degree of dispersion of the global connectivity level results. At the subRCI level, comparison of standard deviation between rounds enables us to ascertain the dispersion in each factor's evolution between the two periods. In both cases (RCI and subRCI), dispersion will relate to the differences in results reached by the considered units, that is, each of the five projects.

The RCI has increased slightly between R1 (45.97%) and R2 (47.79%). In both cases, standard deviation is similar and represents less than a unit (0.99 in R1 and 0.90 in R2), whereupon it can be deduced that the dispersion of results among the projects comprising the global scientific collaboration network is not very high.

Level 4: proposes *mobilizing the socio-technical network towards the field of pertinence identified in the previous levels* (e.g. from Factor 5, *Potentiality*, of the Conceptual Framework) to accompany the inclusion of a variety of actors and perspectives through movements that help to develop socio-technical robustness. Inspired by the Actor-Network Theory and in the RRI vein, network learning is possible during the inclusive, deliberative and projective development process of mutual science-society commitments. We propose the inclusion of digital research field tools, in particular, the *issue crawler*, used to produce cartographies of actors and relations according to subjects and areas (*issue network*).

Level 4 proposes five steps whose objective is to feed the process of problematization of research projects through an interactive process of inclusion of actors and perspectives oriented around the theme-based cartographies acting as mechanisms that propel the construction and deconstruction of “theme areas”:¹

1 – Identification of the current and potential scientific collaboration network. This corresponds to the results obtained with REQUANET Level 1 and reflects a description of the current and potential scientific collaboration network.

2 – Delimitation of the theme area. This is based on the research topic definition, and the overall and specific objectives formally stated in the project. It establishes an initial problematization framework.

3 – Cartography 1 movement. This involves drawing up the problematization network atlas using tools relating to the field of “*digital social research*” in order to identify thematic networks or *issue maps* and the main actors concerned.

4- Cartography 2 movement. This mobilizes reflexive and projective contrast by summoning the actors identified in step 3. It involves mobilizing their *concerns* and priorities and questioning the limits of the *unanalyzed* and the *certitudes*. The dual effect of configuration and reconfiguration of the problematic situation is achieved and, therefore, of the theme area, which is initially delimited with priority, from the cognitive viewpoint (step 2).

Step 5 - Cartography 3 movement. This signifies that network learning has allowed the issue map dynamic to be proposed from an *open, recursive* and *participative* viewpoint (Venturini 2012). The problematization network remains active in the reflexive commitment surrounding the reconfigured thematic area and reconfiguration, in turn, modifies the problematization network’s composition.

This level is still in the process of methodological development and has not been applied to this case study.

¹ This proposal is partially inspired by Venturini’s cartography of controversies (Venturini, 2010, 2012), and the experiences analysed in the MACOSPOL project: <http://www.mappingcontroversies.net/>

SUMMARY

REQUANET is a methodological proposal to operationalize the RRI approach. The development of four levels allows the socio-technical network to be strengthened through inclusive and reflexive movements that configure mutual and sustained commitments among the actors. These movements improve the levels of socio-technical integration that, according to our conceptualization, result in better conditions of relational quality.

REFERENCES

- EC (2015), *Indicators for promoting and monitoring Responsible Research and Innovation. Report from the Expert Group on Policy Indicators for RRI*, European Commission, Directorate-General for Research and Innovation, Brussels.
- Owen, R., Macnaghten, P. & Stilgoe, J. (2012) 'Responsible research and innovation: From science in society to science for society, with society', *Science and Public Policy*, 39 (6), 751-760.
- Venturini, T. (2010) "Diving in magma: How to explore controversies with actor-network theory", *Public understanding of science*, 19 (3), 258-273.
- Venturini, T. (2012) 'Building on faults: how to represent controversies with digital methods', *Public Understanding of Science*, 21 (7), 796-812.
- von Schomberg, R. (2015). "From the responsible development of new technologies towards Responsible Innovation". In J. Britt Holbrook and Carl Mitcham (eds.). *Ethics, Science, Technology, and Engineering: A Global Resource (ESTE2)*, MacMillan Reference USA, Gale.
- Wickson, F., & Carew, A. L. (2014). 'Quality criteria and indicators for responsible research and innovation: learning from transdisciplinarity', *Journal of Responsible Innovation*, 1(3), 254-273.



What knowledge counts? Insights from an action research project using participatory video with grassroots innovation experiences ¹

Alejandra Boni*, Monique Leivas, Alba Talón, Teresa De La Fuente, Victoria Pellicer-Sifres, Sergio Belda-Miquel, Aurora López-Fogués and Begoña Arias

*aboni@ingenio.upv.es

INGENIO, CSIC-UPV, CPI, Camino Vera s/n, Valencia, 46022, (Spain)

Emails in order of co-authorship: moniqueleivas@hotmail.com; altavialba@gmail.com; teresadelafuenteespinosa@gmail.com; vicpelsi@ingenio.upv.es; serbelmi@ingenio.upv.es; aulofo@gmail.com; mariasga@hotmail.com

All co-authors are researchers or associates of INGENIO (CSIC-UPV).

ABSTRACT

This paper presents a contribution on a participatory action-research process using Participatory Video (PV) methodology. During six months, a group of 6 facilitators and 9 members of two grassroots innovation initiatives (Solar Dómada and Fuel Poverty Group) took part of the process and produced two videos during a five-stage PV process, from initial definition and planning to public screening and debate of the videos. We present some insights from that research using an original framework developed to analyze PV process: the eParc Cube. This framework examines the interaction between knowledge production, participation and communicative spaces that happen during PV. We conclude reflecting on the social relevance of that kind of research considering the impact among of the co-researchers of both process and products.

INTRODUCTION

How do we know that the knowledge produced through research has a social impact? Which impact are we achieving? Who defines them? All these issues are of particular relevance in this Conference dedicated to explore the peripheries in the production and measurement of scientific knowledge.

Through this research-in-progress we show some insights of a recent action-research process using participatory video (PV) as a tool. This research has been conducted from October 2015 to March 2016. A group of 6 facilitators (co-authors of this paper) and 9 members of two grassroots innovation initiatives (one of them also co-author of this paper) took part in this research and produced two videos during a five-stage PV process, from initial definition and planning to public screening and debate of the videos. The two initiatives pointed at bottom-up, social, alternative and empowering production of energy and of space.

¹ This work was part of the Project Nuevas perspectivas para repensar el cambio climático desde la innovación social de base. Abordaje desde el desarrollo humano, el aprendizaje y la ciudadanía (CSO2013-41985-R) granted by the Spanish Ministry of Economic and Competitiveness. It was also supported by the Centro de Cooperació al Desenvolupament de la Universitat Politècnica de València. We are grateful to Gynna Millán for having prepared the two figures of this paper and for her comments and editing.

As the main topic of this conference recalls, the knowledge that has been produced could be considered “peripheral” for political, academic and geographical reasons. Firstly, because it is knowledge produced from social groups, which have in common their activism for a change of the mainstreaming model of development; secondly, because it has been produced through a “peripheral” research methodology in academia, such as action-research; thirdly, because the research has been carried out in Valencia, a city considered on the outskirts of the scientific knowledge production predominant today.

In section 2 we describe the methodology used and the main characteristics of the two organizations involved; in section 3, we explain the analytical framework used (the eParcube, Boni and Walker, 2016) to collect the evidences; in section 4 we describe some results and we conclude with some insights on the social relevance of this research and how has been captured and measured.

CASE STUDY. A PV PROCESS WITH TWO LOCAL GRASSROOTS INNOVATION INITIATIVES

PV has been largely used as a method and a process with the aim of empowering individuals and communities through sharing stories and making videos depicting their own realities, challenges and aspirations for the future (White, 2003). PV can be considered as one of the many manifestations of the relationship between media and development (Scott, 2014) and also as a tool under the umbrella of participatory action methodologies.

PV is a wide field, which allows a wide range of approaches and perspectives (High et al, 2012): some use it as a method for research (Oliver et al, 2012), while others regard it as a tool and a process to foster awareness for local communities (White, 2003; Plush, 2012). Other authors have explored it as a way to influence policy making (Wheeler, 2012), although, in the same experience, a PV process could aim to achieve more than one of those goals. According to Shaw (2013) there is neither a single nor correct method to approach a PV process and what happens in each experience is very contextual and could lead to very different outcomes.

In our particular case, PV has been used as a research method to try to grasp the contextual knowledge produced and as a way to empower members of local initiatives through different cycles of reflection and action. Also, to produce an output (the two videos²) that can be useful for the goals of the different participants, for instance, for the local organisations as a tool to show and disseminate their activities and add new constituencies and for the group of facilitators, as a way to showing a peripheral way of conducting research and discuss the social relevance of it.

² Available at <https://repensandoelcambioclimatico.wordpress.com/5o-ciclo-proyeccion-publica/>

Participants

The two local organizations were Solar Dómada (<http://domonmada.blogspot.com.es/>) and Fuel Poverty Group (hereafter FPG) of the “*Plataforma por un Nuevo Modelo Energético*” or “Platform for a New Energetic Model” (hereafter Plataform) (<http://www.nuevomodeloenergetico.org/pgs2/>). As we will see in the brief descriptions below, both organizations can be understood as grassroots innovations (GI) which, according to Seyfang and Smith (2007: 585), can be defined as:

"networks of activists and organizations generating novel bottom-up solutions for sustainable development; solutions that respond to the local situation and the interests and values of the communities involved".

The first group is the Solar Dómada, a group of people who are occupying a private plot, highly deteriorated at the time of their occupation (2013), as a way to assert the need for social spaces in the neighbourhood. Solar Dómada also seek to highlight that another kind of coexistence between neighbours is possible; one based on respect and intercultural coexistence. In the centre of the plot is the Garden of Ca Favara, one of the symbols of neighbourhood participation, involving more sustainable practices of food production and consumption.

The second GI is the Fuel Poverty Group, a very new group of volunteers, mainly university students that want to challenge fuel povertyⁱ by giving advice on how to reduce fuel consumption. This group is part of a wider network named the Platform for a New Energy Model, which works towards a more democratic and sustainable energy model.

The two groups have a common aim behind their activism in that they both seek a more equitable, democratic and sustainable livelihood. The differences between them lie in: the area in which they are located (energy and production of urban space); the age and characteristics of their members (university students in the case of Fuel Poverty Group and people of different ages, educational levels and careers in the case of Solar Dómada); and their strategies (information and technical advice in the case of Fuel Poverty Group and occupation of urban space in the case of the Solar).

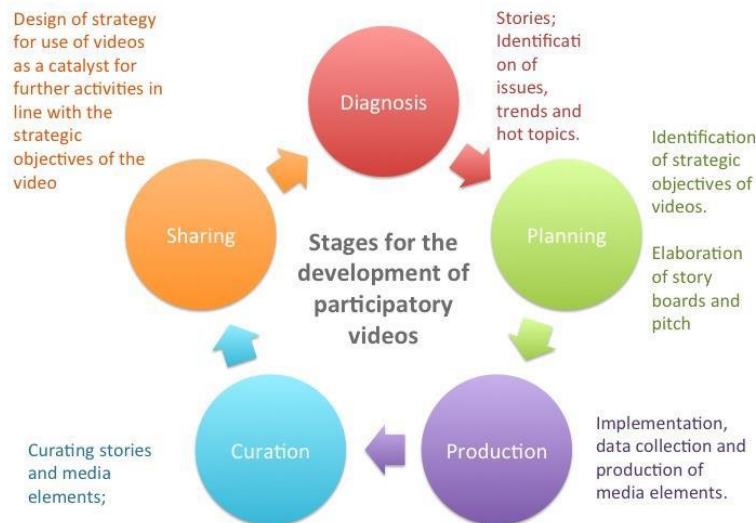
The other participants in the process were the facilitators, all researchers and collaborators at INGENIO, a Spanish institute devoted to knowledge management and innovation (<http://www.ingenio.upv.es/en>).

PV stages

Figure 1 depicts and explains the five phases of the PV process

Figure 1: Stages of Participatory Video Development.

Source: Millán and Frediani 2014



In the first phase, diagnosis, participants identified the most relevant issues. This phase occurred in two types of communicative spaces: within each of the two groups and between the two groups and with the facilitators. In the case of Solar Dómada, the intra-group space was particularly important as it enabled a reconstruction of the history of the group. In the case of the Fuel Poverty Group, the interaction between this group and the other participants enabled them to think about the narrative of the video, embracing a broader perspective of fuel poverty.

The second phase was planning, where the storyboard was developed. This occurred primarily inside groups and then it was socialized in a communicative space of a collective nature, which was also very much appreciated by the participants, allowing them to reflect on the narratives and contents of the two videos.

The third phase was the video production. In the case of the Fuel Poverty Group, the participants asked people outside the action research about the significance of fuel poverty or how they felt about being labelled “energy poor”. As we will discuss in section 4, these interactions were a very important source of learning about rethinking the idea of fuel poverty and the scope of performing energy consultancy as a mechanism to deal with it. In the case of Solar Dómada, the production phase stimulated a variety of communicative spaces between group members and the neighbourhood, providing various perspectives on the plot. The contribution of the facilitators at this stage was to provide technical assistance in recording.

The fourth phase was the publication (curation) of the two videos, which in the case of Solar Dómada was conducted with the help of an external facilitator, while in the case of the Fuel Poverty Group, the task was taken on by the group itself. There was a collective communication space where videos were pre-viewed internally. For the Fuel Poverty Group, this space of collective discussion allowed them to refine the video narrative.

The 8-minute video by Solar Dómada (available at <https://www.youtube.com/watch?v=FUUMTSxU6Iw>) presents the occupied plot as a place where coexistence between neighbourhoods is promoted and more sustainable lifestyles are demanded, which are respectful of the differences between cultures and between generations. At the heart of the plot there is a small orchard, literally dug into the cement, symbolizing a space of resistance against a model of the unsustainable and individualistic city in a peripheral and difficult urban environment. The second video (6'50"), available at <https://www.youtube.com/watch?v=Ke6fQxCrnro>, illustrates a recent problem in the Spanish context; that of fuel poverty. The video shows evidence of what is meant by fuel poverty and how conducting a review of the entire energy consumption of a household can lead to improved energy efficiency.

The PV cycle ended with the public presentation of the videos in an emblematic site in the city of Valencia, due to its political character (*Ca Revolta*). After the screening there was an interesting dialogue between group members, facilitators and the audience, composed of activists and academics and neighbours of the Solar Dómada.

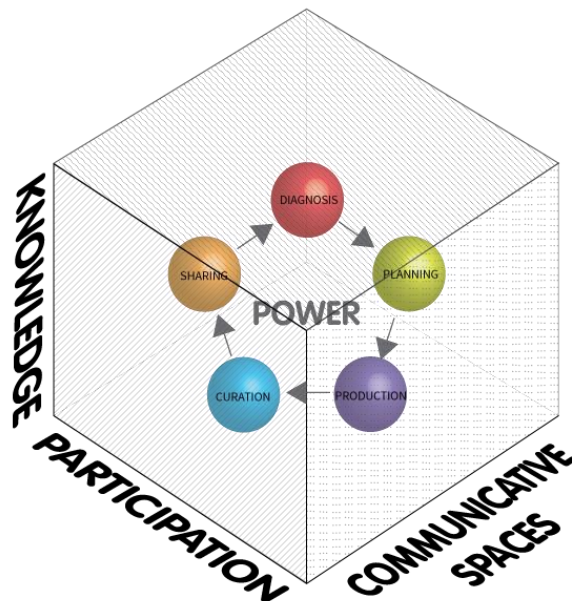
THE EPARC CUBE. A FRAMEWORK FOR THE ANALYSIS OF THE PV PROCESS

To conduct an analysis of this experience we will use an original framework designed to capture the digital participatory action research process. This framework has been developed by Boni and Millán (2016) and was inspired in previous works by Boni and Walker (2016), Frediani (2015) and Gaventa (2006).

The first category for analysis is the idea of communicative spaces that can be understood as forums in which people join as co-participants in the struggle to remake the practices in which they interact (Kemmis and McTaggart, 2005:563). The same authors define practice as real, material, concrete and particular actions of particular people in specific places and can comprehend what people do, how people interact with the world and with the others, what people mean and what they value, the discourses in which people understand and interpret the world (Kemmis and McTaggart, 2005: 565).

The second category is participation. According with Bradbury (2010: 104), participation can be considered in a broad spectrum: from a minimum involvement of practitioners (for example, in a needed consultation) to having those practitioners as co-researchers and co-designers.

The last category is knowledge; through participation in communicative spaces knowledge is produced, assumed not only as an understanding of the topics addressed, but also practical knowledge (the skills developed) and the values that underpin the knowledge produced (Kemmis and McTaggart, 2005: 565). Each of these three elements – communicative spaces, participation and knowledge – will form the axes of a three-dimensional figure, a cube. The use of the cube aims at visualising complex interactions among dimensions in the analysis of participatory processes. In our case, it aims at representing the intersections that occur between knowledge, power and participation within communicative spaces, taking place during the cycles of reflection and action in the phases of the PV. For our analytical understanding, in Figure 2, we find the figure that represents the PV process (see fig.1), inside the cube. In the interactions between the three axes, issues of power emerge and shape the kind of participation and knowledge produced (Gaventa and Cornwall, 2008).

Figure 2: The ePARC framework (Boni and Millán, 2016)

EVIDENCE

We will base our evidence on participant observation conducted throughout the process along with three group interviews to members of the two GIs at the end of the PV process. In the case of Solar Dómada, two group interviews were conducted: the first with 3 women participating in the GI and the second with two men. The reason for doing it this way is that, during the PV, a difficult power relation between one of the women and the two men was detected. To enable the interview to flow more naturally, it was decided to separate the two groups. In the case of the Fuel Poverty Group a single group interview was conducted with a woman and a man. All the interviewees give us their informed consent.

We will begin this section by analysing the potential of communicative spaces (both collective and within the groups) to create knowledge and foster participation. Starting with the collective communicative spaces mentioned before, both groups acknowledge that the first collective meeting was highly motivating and exciting. As noted by one of the members of the Solar Dómada:

“It was very encouraging to see that your team [INGENIO team] was interested in our initiatives and because the problems we often have is making ourselves understood by our neighbours... I thought it was a good opportunity to become known in the neighbourhood... also to try something new, editing a video is far from what we normally do.”

In terms of the knowledge produced, we can identify the second collective moment that happened at the planning phase as being extremely powerful – when the two organizations shared storyboards. During moments of dialogue, participants were able to contrast their visions on the themes that would be address in the videos. For example, one of the members of the Fuel Poverty Group indicated:

“XX told us that energy is not only electricity... there is solar thermal energy in the roofs of the houses... [all of these] are reflections from other points of view that you can get if you talk to people, and especially if you talk to groups that are already committed... [this is] where richness lies”

In the case of the Fuel Poverty Group, the collective moment helped the group to adopt a less paternalistic perspective of fuel poverty. Their first option was to show one person affected by fuel poverty and how the energy consultancy could help to reduce her energy expenditure. After the ideas exchange during the second collective moment, the group decided to include a more political perspective of fuel poverty, introducing references in the storyboard concerning the energy oligopoly that exists in Spain and which hinders better energy consumption.

Regarding facilitation, members of the two groups expressed that horizontal relations between facilitators and members of the two GIs had a positive and significant impact on communication and exchange of ideas. It was also highlighted in the final meeting that collective spaces had been planned and managed in a very careful way. They were experienced as pleasant and friendly spaces, where people felt comfortable and relaxed, having a positive effect on people's participation. The importance of the emotional aspect in the process has been one of the greatest learning aspects for the facilitation team. Relationships between people are crossed by emotions, and creating communicative spaces where these emotions can be channelled positively is essential in order to generate more knowledge sharing and enhance participation.

With regard to the communicative spaces that have occurred within groups, for Solar Domada's members, the exchanges that happened during the diagnostic and planning phases were very important to reconstruct the history of the organization and the role played by each of its members. As noted by one of the participants:

“We remember especially when we were recalling those moments with pictures... they were very emotional moments... I loved it when all of us answered without digressions what we wanted to show in the video... we had never seen such an organized and respectful relationship as the one that occurred that day”

In the case of the Fuel Poverty Group, one of the most interesting communicative spaces from the perspective of knowledge production took place at the production stage when interviewing a woman affected by fuel poverty. The interviewer noted that the most shocking thing was to realize that the woman wouldn't have considered herself fuel-poor if she had been living on her own but she would reconsider this position if this affected her ability to meet the basic needs of her family.

Another important learning aspect for the members of the GIs was the limitations of their voluntary action as a way to challenge fuel poverty. As one of them indicated:

“The difficult part is that we can help reducing the bill but we can't help you to get reconnected to the power supply... it's an economic issue... this where we say: we can only go so far as fuel poverty volunteers...”

Finally, the act of making the two videos has also contributed to the acquisition of new technical skills. At the beginning of the process, some of the participants believed they were totally incapable of making a video.

A special mention must be made regarding the power relations that occurred throughout the process; on one side, although the PV process puts the team of facilitators in a position of superiority due to their mastery of the audiovisual tool (Millán & Boni, 2016), this was not a hindrance throughout the process. In the case of the Fuel Poverty Group, the group requested technical support when needed, but much of the technical work was done by the group itself. Participants recognized that the video could have had a higher technical quality, but their attitude was that this was a first approach to the tool, which would enable them to make more videos in the future.

On the other hand, in the Sólár Dómada group, power relations played an important role. In fact, one of the external facilitators ended up in charge of technical tasks and coordination of the PV process, precisely as a way to mediate between group members. This was viewed positively by most participants, because it was the way to “save” the process and finalize the video. However, one of the participants said he would have liked to have more control over the process, but the difficult relations inside the group favoured the delegation of coordination and technical tasks to an outsider.

CONCLUDING THOUGHTS

Although the research has not been completed, and we may need to analyse other collective moments such as the public screening and the potential impact of other communicative spaces produced by the dissemination of the two videos, we can point out some preliminary ideas about the social relevance of a research of this kind.

From the evidence collected, it can be said that the action-research process has had a social impact in terms of knowledge production. It has helped participants to reflect and to problematize the way they understand their “practices”, in the sense proposed by Kemmis and McTaggart (2005). In the case of the Fuel Poverty Group, it has served to rethink their comprehension of fuel poverty and the scope of volunteering practice. In the case of Solar Dómada, it has helped to reconstruct its history as a group and to reflect on the aims of occupied spaces within their neighbourhood. For the team of facilitators, it has also served to rethink the role of facilitation and the importance of the emotional aspect in these processes. Moreover, the VP process had equipped the participants with new technical skills to produce videos and tell powerful stories that could have a social impact using audio-visual language.

With regard to the outputs of this PV process, the two videos are extremely meaningful for the two organizations, as they constitute another communication tool that could help in the diffusion of their social causes they stand for.

We argue that the evidence presented can also be considered indicators of social relevance. Certainly, this is a small scale research and, from a cost-benefit perspective, this could be considered too expensive and time consuming. But which criteria must prevail in measuring the social impact of a research? Could the creation of a contextual, participatory and transformative knowledge be considered a significant indicator to measure the social relevance of research? All those issues should be discussed and problematized in conferences like this one.

REFERENCES

- Boni, A. & M. Walker. *Global Human Development: Theoretical and Empirical Insights for Social Change*. London: Routledge, 2016.
- Boni, A. & Millán, G. (2016). Action-Research using Participatory Video. A learning experience in San Lorenzo, Castellón, Spain. *Commons. Revista de Comunicación y Ciudadanía Digital*, Vol. 5 N. 1 pp. 113-135.
- Frediani, A. A., “Participatory Capabilities in Development Practice.” *DPU Working Paper* 178 (2015). Accessed September 28, 2015.
www.bartlett.ucl.ac.uk/dpu/latest/publications/dpu-working-papers/WP178.pdf
- Gaventa, J. “Finding Spaces for Change. A Power Analysis.” *IDS Bulletin*, 37(6): 23-33. 2006.
- Gaventa, J. & A. Cornwall. “Power and Knowledge.” In *Sage Handbook of Action Research: Participative Inquiry and Practice*, 2nd edition, edited by P. Reason and H Bradbury. London: Sage, 2008, pp. 71-81.
- High, C., Singh, N., Petheram, L., & Nemes, G. (2012). Defining Participatory Video from Practice. In E. J. Milne, C., Mitchell, & N. De Lange (Eds.), *Handbook of participatory video*. (pp. 35-48). Lanham, Md: AltaMira Press.
- Kemmis, S. & McTaggart (2005) Participatory Action Research: Communicative Action and the Public Sphere. In Denzin, Norman K. and Lincoln, Yvona S., (eds.) *The SAGE Handbook of Qualitative Research*. 3rd ed. London: SAGE Publications
- Mitchell, & De Lange, N. (Eds.) (2012), *Handbook of participatory video*. (pp. 365-382). Lanham, Md: AltaMira Press.
- Millan, G.F., & Frediani, A.A. (2014). *Terms of Reference of the Summer School*. Castellón: UJI.
- Olivier, T., de Lange, N., Creswell, J.W., & Wood, L. (2012). Mixed method research in participatory video. In E. J. Milne, C., Mitchell, & N. De Lange (Eds.), *Handbook of participatory video*. (pp. 131-148). Lanham, Md: AltaMira Press.
- Plush, T. (2012). Fostering Social Change through Participatory Video: A Conceptual Framework. In E. J. Milne, C., Mitchell, & N. De Lange (Eds.), *Handbook of participatory video*. (pp. 67-84). Lanham, Md: AltaMira Press.
- Scott, M. (2014). *Media and Development*. London: Zed Books.
- Shaw, J. (2013). *Using Participatory Video for Action Research*. Real Time. (real-time.org.uk)
- Seyfang, G., & A. Smith. (2007). “Grassroots Innovations for Sustainable Development: Towards a New Research and Policy Agenda.” *Environmental Politics* 16 (4): 584–603. doi:10.1080/09644010701419121.
- Wheeler, J. (2012). Using Participatory Video to Engage in Policy Processes: Representation, Power, and Knowledge in Public Screenings. In E. J. Milne, C.,
- White, S. A. (2003). *Participatory video: Images that transform and empower*. New Delhi: Sage.

ⁱ Fuel poverty can be understood as the difficult situation faced by a household that can't afford to pay for their energy consumption, leading to a lack of normalized access to gas and electricity. This situation brings about a deprivation of the freedom to live a decent life (Pellicer & Lillo, 2014).



A proposal for measurement of science and innovation culture¹

Asako Okamura²

*a-okamura@grips.ac.jp

SciREX Center, National Graduate Institute for Policy Studies (GRIPS) 7-22-1, Roppongi, Minato-ku, Tokyo, 106-8677, JAPAN

ABSTRACT

Why do perceptions about the negative and positive aspects of science, technology, and innovation differ among individuals and across countries? What types of technology do we fear and what types do we embrace? Amongst the general population, which group is most comfortable with new technology and which group is most sceptical about its diffusion? Why are scientific careers popular in some countries and not in others? In the end, is there any relationship between appreciation for science and well-being? How is our relationship with technology linked to national competence and national innovation systems?

These questions are of particular importance for science, technology, and innovation policy these days, as shown in some increasingly used policy concepts and keywords, such as ‘responsible research and innovation’, ‘societal impact of science’, ‘science and society’, and ‘innovation for societal issues’. As science and innovation activities are globalized, these ‘cultural’ factors have also gained global importance.

In light of the importance of science and innovation culture as a foundation of science, technology and innovation policymaking, a future research agenda to advance our understanding and measurement is proposed.

INTRODUCTION: BACKGROUND AND OBJECTIVE OF THE STUDY

In the fast-paced modern world, science and technology impact on all areas of life, at both an individual and organisational level. However, the degree of acceptance of science and technology varies across individuals and organisations both within countries and at a cross-cultural level.

Why do perceptions about the negative and positive aspects of science, technology, and innovation differ among individuals and across countries? What types of technology do we fear and what types do we embrace? How much detail do we need to know regarding the benefits and harmful effects of emerging technologies? Amongst the general population, which group is most comfortable with new technology and which group is most sceptical about its diffusion? Why are scientific careers/entrepreneurs popular in some countries and not in others? How do we know the level of innovativeness of citizens? In the end, is there any relationship between appreciation for science and well-being? How is our relationship with technology linked to national competence and national innovation systems?

These questions are of particular importance for science, technology, and innovation policy these days, as shown in some increasingly used policy concepts and keywords, such as

¹ This is a preliminary draft. Please do not cite without the author’s permission.

² The author greatly thanks the former colleagues at the Organisation for Economic Co-operation and Development (OECD) as well as colleagues at the SciREX center. The views and opinions in this paper are those of the author, and not necessarily those of the organizations.

‘responsible research and innovation’, ‘societal impact of science’, ‘science for society’, and ‘innovation for societal issues’.

As OECD (2014) suggests that ‘innovation is influenced by certain social and cultural values, norms, attitudes and behaviours which may be described as an “innovation culture”,’ there may be ‘science and innovation culture’ as underlying conditions for national science and innovation system.

However, the question is how we measure these concepts and acquire meaningful indicators to inform policy-making. Traditionally, these questions have been partly explored in the field of Science and Technology Studies (STS) (See Bauer (2008) for a survey). Traditional frameworks may need to be updated, to capture extended link between science and innovation.

The objective of this paper is to define a concept of ‘science and innovation culture’, as a working definition. Then it proposes some ideas for advancing measurements and constructing indicators for capturing ‘science and innovation culture,’ by briefly reviewing previous studies. Then it concludes by offering a future research agenda to advance the relevant studies.

DEFINING A ‘SCIENCE AND INNOVATION CULTURE’

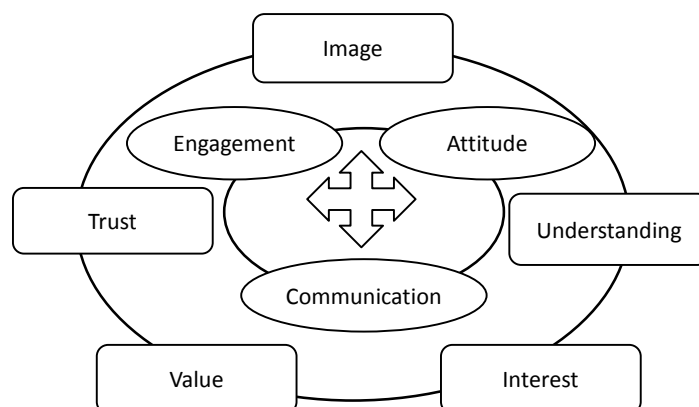
To clarify the scope of this paper, we first need to provide a definition of ‘science and innovation culture.’ However, no stylized definition is yet available. The concept of ‘science and innovation culture’ has not been discussed explicitly in previous studies with few exceptions and we do not yet have sufficient evidence to validate this definition.

The OECD (2014) states that ‘innovation is influenced by certain social and cultural values, norms, attitudes and behaviours which may be described as an “innovation culture”.’ Godin (2013), meanwhile, states that:

A culture of science is a culture defined, partly or wholly, by and through science. It is a culture in which a central set of institutions and activities are concerned with science, broadly defined, including the users and effects of the productions arising from these institutions and activities. ... a culture of science is that sum of dimensions or subsystems that includes: Institutions (research); Productions (graduates, knowledge, technologies); Diffusion, use and users (education, transfer, communication); Impacts (effects on society, the economy, the individual); and Environment (laws, economic system, social values).

We propose defining ‘science and innovation culture’ as the basis of individual and collective values, choices, behaviours, and risk-preferences related to science and innovation (both in our daily lives and in the longer term, including career choices), and a source of knowledge creation. These eventually affect the economy and society in which we live through the consumption pattern of science and innovation and the accumulation of human capital. This serves as a core foundation of a national innovation system, which, in turn, affects how the science and innovation systems work, how scientific research is conducted, the creation of science policy, and the way of democracy.

In our definition, ‘science and innovation culture’ refers to both individual level culture and institutional level culture, as in universities and research organizations. It is also categorized as multi-dimensional, including the dimensions of individual perception, attitudes, acceptance, trust, behaviour, utilization, rejection, concerns, incentives, openness, skills, capacity, and career paths. The key elements may be summarised in the below diagram.



‘Science and innovation culture’ also concerns the relationship between the general public (citizens, individuals), research community, the business and industrial sector, and science policy making.

The main difference between the proposed definition and the two previous ones lies in its inclusion of innovation into science culture. This is particularly important as ‘science and technology’ and ‘innovation’ have become increasingly convergent, though its differences in natures should be noted.

MEASUREMENT OF ‘SCIENCE AND INNOVATION CULTURE’

The next objective of this paper is to propose some ideas for advancing measurements and constructing indicators for capturing ‘science and innovation culture.’ Currently, it is virtually impossible to untangle and visualize the complex interactions between these elements and gain a fresh perspective on how relevant policies should be formulated and evaluated. We first review methodologies in previous studies, briefly.

The study of the link between science and society is commonly described in the literature as Public Understanding of Science (PUS), Public Perceptions of Science (PPS) and/or Public Engagement with Science (and also technology) (PES), each term emphasizing a particular dimension of the link.

A variety of methodologies is used, differing in objectives, content and collection instrument used, to measure the link between science and society. Bauer and Howard (2013) provide a list of relevant publications within the journal *Public Understanding of Science* from 1992 to 2011. Methodologies used in publications are distributed as follows: Survey (113 publications); Case study (112); Content analysis (99); Critical analysis (58); Discourse / rhetorical analysis (29); Document analysis (28); Experiment (8); Film/TV analysis (20); Interviews (56); Observation (17); Other (29).

Evidence on this topic is typically collected from surveys of citizens, drawn from the general population, although other quantitative and qualitative methods are being applied such as media content analysis, focus groups, and in-depth ethnographic studies. Furthermore, actors in the science system – such as scientists and decision makers in government or industry– can also be effective sources of information on the direct linkages between science and society.

A number of surveys have been conducted at international level, as well as at national level, including EU barometers (EC, 2014), etc. NSB (2016) synthesises national and international

efforts. The Ibero-American Network on Science and Technology Indicators (RICYT) has developed “Manual de Antigua” in 2015, a set of guidelines on measuring public perceptions of science and technology based on the experience of surveys carried out in the region. The manual deals with several dimensions including, interest, information and public engagement; the institutional dimension of S&T; personal appropriation of S&T; attitudes and values to S&T; and socio-demographic and contextual classification.

The public also plays more active roles in science and innovation activities, i.e., citizen science and user innovation. Tracking ‘user-led innovation’ and ‘citizen science’ is another strand of studies to capture ‘Science and innovation culture’. Increased access to data, in this age of open science and innovation, is expected to accelerate the public participation (OECD, 2014a). Jong and Hippel (2013) point that user firms and individual end users, who primarily innovate to satisfy their needs, rather than to sell a product on the market, become the second and increasingly important innovation model revolvers.

In response to its policy concept, responsible research, and innovation (RRI), the European Commission has been developing indicators to monitor RRI, combining performance and perception indicators (EC, 2015).

To validate the definition of ‘science and innovation culture’, we need to understand the role of culture in national science and innovation system. Such analysis might be possible to link perception indicators with main stream indicators on R&D activities and also organizational level culture. Such studies rarely conducted.

The micro analysis is necessary to understand the complex interaction of elements of ‘science and innovation culture.’ Complementary use of other methodologies, including contents analysis, is highly expected especially gauging online dialogue concerning science and technology.

We propose constructing a measurement agenda, by bringing in user perspectives. A future research topic is proposed as below:

- Measurement of dynamics of science and innovation culture: how do individuals form attitudes and trust?
- Measurement of how science and innovation culture influence science and innovation activities
- The connection between science and innovation culture and science and innovation activities (e.g. R&D, Number of researchers, Number of Papers, Patents).
- Measurement of institutional factors of science and innovation culture
- Understanding the gaps between scientists and the public
- Public acceptance of emerging technologies: e.g. AI, Nanotechnologies
- Understanding the public via a public segmentation model
- Understanding micro-mechanisms rather than macro trends.
- Determining public perceptions using micro data (e.g. gender, scientific background)
- Link with subjective well-being (using microdata) (See OECD (2015))
- Determining communication gaps and correspondences using content analysis of social communication media
- Exploring how science excellency influences behaviours and career decisions
- Exploring the role of individuals in driving innovation; science and technology activities

CONCLUDING REMARKS

To advance our understanding and measurements, we propose a forward-looking strategy as below. We offer an incremental approach towards developing a systematic and international comparable measurement framework.

- The first step to facilitate such research is to construct a comparable international database that may serve as a research infrastructure. The potential benefits of using microdata are large, in providing a better understanding of complex phenomena of science and society. Building a microdata inventory as an analytical basis and making it accessible to researchers would enhance academic and policy studies.
- Given the subjective nature of the inquiry, the complimentary use of qualitative methodologies (e.g. case studies, workshops, and focus groups), beyond public opinion surveys, is also required to understand the contexts. New methodologies using internet-based data should also be explored.
- Identifying and articulating user needs are necessary to shape a measurement scope. Dialogue with policy communities is an important step towards determining the core questions that need to be asked.

Organization of Measurement Manual

Finally, we propose making a measurement guideline on the subject of science and innovation culture by an international coordinated effort. This is important, not only in itself but a process to make it matters in community development and advances in measurement. We conclude by proposing a draft structure as below:

- Background and the needs for a measurement manual
- Scope of the measurement manual
- Unit of analysis
 - Individuals
 - Organizations (e.g. Research institutions, industry-university cooperation agencies, Civil Society)
- Methodologies
 - Standardized sampling design
 - Harmonised questionnaire
- Survey items:
 - Individual (citizen): Public understanding, interest, values, trust, image, acceptance, communication, participation, and involvement, regarding science, technology, and innovation. Entrepreneur spirits and risk preferences.
 - Organization: Scheme for connecting science, innovation, and society
 - Linkage between organizations and individuals
- Measurement issues
- New methodologies
- Using big data (e.g. web-based tools)
 - Possibilities
 - Measurement issues

REFERENCES

- Bauer, M. (2008). “Survey research on public understanding of science,” in M. Bucchi and B. Trench (eds) *Handbook of public communication of science and technology*, Routledge, London/New York.
- Bauer, M., and S. Howard (2013), “Public Understanding of Science: compiled bibliography, 1992-2011”, London, available at http://pus.sagepub.com/site/misc/PUS_book_v6_AG.pdf.
- European Commission (2014), “Special Eurobarometer 419: Public Perceptions of science, research and innovation”, http://ec.europa.eu/public_opinion/archives/ebs/ebs_419_en.pdf
- European Commission (2015), “Indicators for promoting and monitoring Responsible Research and Innovation”, Report from the Expert Group on Policy Indicators for Responsible Research and Innovation.
- Godin, B. (2013). “The culture of science and the politics of numbers”, in Bauer, M., Shukla, R., and Allum, N. (eds) *The culture of science – How the public relates to science across the globe*, Routledge, New York.
- National Science Board (2016), Chapter 7, “Science and Technology: Public Attitudes and Understanding”, Science and Engineering Indicators 2016, available at <http://www.nsf.gov/statistics/seind16/content/chapter-7/c07.pdf>.
- OECD (2014). “Building a science and innovation culture”, in *OECD Science, Technology and Industry Outlook 2014*, OECD Publishing. Available at http://dx.doi.org/10.1787/sti_outlook-2014-33-en
- OECD (2015). *OECD Science, Technology and Industry Scoreboard 2015: Innovation for growth and society*, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/sti_scoreboard-2015-en



Scientific culture in Colombia. A proposal of an indicator system for science technology and innovation¹

Clara Inés Pardo Martínez* and William H. Alfonso P.**

*cipmusa@yahoo.com

Colombian Observatory of Science and Technology, Carrera 37 No. 15-39, Bogotá, Colombia

**William.alfonso@urosario.edu.co

University of El Rosario, Calle 12C No. 6-25, Bogotá, Colombia

ABSTRACT

In last decades, scientific culture has become a key element of Governance of Science, Technology and Innovation in the countries where it is important to determine measurement to analysis trends on scientific culture. Research questions that guide this paper are the following: i. What are information needs on scientific culture in Colombia?; ii. How can be measured scientific culture?; iii. What is the adequate structure for indicators of scientific culture?. In order to answer these questions, a mix of methodologies is used. First, we review the literature on scientific culture and indicators related to this topic. Second, we made a series of interviews with staff members of Colciencias to determine requirements of measurement on scientific culture. Third, with this information, we built an information matrix to prioritise information and determine indicators with respective metrics, and sources according to relevance and cost-effectiveness of estimation. Fourth, from indicators formulated and an indicator system is proposed determining for every dimension of scientific culture indicators related to inputs, process, and outputs designed indicator sheets that includes definition, objective, sources aggregation levels, time series, and calculation methods for indicators proposed. This study achieves formulate an indicator system from the definition of scientific culture a and its dimension proposing around 60 indicators through a multidimensional model that integrates different elements of scientific culture such as the individual and society establishing indicators to measure inputs, process and outputs in general form and specific initiatives for Colciencias.

INTRODUCTION

In last decades, scientific culture has become a key element of Governance of Science, Technology and Innovation in the countries where it is important to determine measurement to analysis trends on scientific culture.

National agencies of science, technology and innovation around the world have sought to generate strategies and to support initiatives with the aim to strength and reveal articulations between science and society through diverse activities aimed at non-expert publics. These activities have been denominated of multiple shapes such as divulgation, popularization, public communication of science and technology, science journalism, or the social appropriation of science and technology. The majority of these requirements are based on the

¹ This work was supported by Colciencias and Colombian Observatory of Science and Technology (OCyT)

assumption that there is a broad gap between “the sphere of science” and “the society” that it is necessary to close through initiatives that allow that these two elements achieve a better articulation especially in knowledge-based society.

Around the world, scientific culture has been defined from different perspectives. However, the definitions have common element related to the social appropriation of science and technology that implies financial, regulatory, coordinating, educational, or communications guideline. These elements and activities conform the collective part of scientific culture (Godin and Gingras, 2000).

In Colombia, scientific culture has been defined by Colciencias (2016) as the promotion and strengthening of science, technology and innovation culture understood as a set of beliefs, attitudes, practices and values of citizens that have been permeated by science and technology (S&T), promoting:

- a. Possibilities to claim the status of S&T in the Colombian society.
- b. Interest, enthusiasm, the reinforcement skills and re-signification of S&T as the choice of life.
- c. Participation of citizens of consistent and reflexive form in democratic scenarios on scope, limitations and risk of S&T.
- d. Insertion of scientific and technologic knowledge for solution of social, environmental and productive problems.
- e. Generation and promotion of innovative and entrepreneurial ideas from science and technology.

Moreover, to develop this definition, it has established three key dimensions: Training of human resource for science, technology and innovation, public communication of science, and social appropriation of science, that are defined as follows by Colciencias:

Training of human resource for science, technology and innovation is a set of attitudinal (way of acting), ability (capacities) and psycho-affective (context conditions) factors that demonstrate the inclination of a person to science, generation or transformation of knowledge. These factors identify or develop early or late from experiences with scientific objects and processes, that are expressed in scenarios of science and society over the life and the different training levels.

Public communication of science, technology and innovation is a wider process concerned with the transfer of knowledge from one subject or group of subjects to another. Moreover, knowledge as being transferable without important modifications from one context to another, so that it is possible to take an idea or result from the scientific community and bring it to the general public (Bucchi, 2008). The communication process from researchers to popular science thus be exemplified as like a funnel that subtracts refinements and shades of significance from the knowledge that passes through it, reducing it to simple facts recognised with certainty and incontrovertibility. Fleck stresses that this progressive solidification of knowledge then exerts an influence on specialists themselves (Bucchi, 1998).

Social appropriation of science, technology and innovation defined by Colciencias (2010) as an understanding and intervention process of relationships between techno-science and society built from active participation of diverse social groups that generate knowledge that it is characterized by a broad concept higher those other similar objectives (such as disclosure,

popularisation, scientific communication, etc.). This process includes: i. actors with diversity of cultures and interests front science and technology subjects; ii. scenarios where scientific ideas are presented as an articulation process between scientific and technological knowledge with diverse forms of wisdoms; iii. Education and communication materialities that are designed and used in these processes with intentions of different actors and approaches; iv. socio-scientific situations that generate interest or matter of concern at different actors; v. Empowering society with science and technology is a concept developed by OECD (2015) emphasizes on the level to which citizens participate in innovative processes, the degree of sophistication of demand, and readiness to accept and recognise the potential of science and technology.

From these concepts on definition of scientific culture and its dimension, it is necessary to develop and identify a set of indicators to measure trends and dynamics of scientific culture in Colombia with the aim to improve policy instruments and strength the different dimensions to achieve a better concept and integration of science, technology and innovation in the Colombian population as a key factor for development and an immigration to knowledge-based society.

In general, scientific culture has been measured through surveys on the public perception of science and technology, which have several limitations that presume a public deficient in knowledge, attitude or trust of population, theory testing, probability and uncertainty, and difficulty to replicate experiments (Collins and Pinch, 1993) that should be compensated contextualizing survey research, defining new cultural indicators, integrating datasets and doing longitudinal analysis, and including other data streams (Bauer, 2007). These elements are key to improve the measurement of scientific culture using other strategies such as indicators.

The main objective of this paper is to design an indicator system to measure science, technology and innovation culture to evaluate and determine its trends and dynamics in Colombia as key tool for decision-making and design of adequate policies and instruments that promote science and technology in the country as development engine.

The specific objectives of this study depart from the following aspects: i. To define indicators related to scientific culture and training of human resource for science, technology and innovation; ii. To determine indicators on appropriation of science and technology; iii. To establish indicators to measure activities of science communication; and iv. To design general indicators on scientific culture. To define and develop these indicators will allow to make monitoring and control of different initiatives of scientific culture in the country taking into account scope, results, effects and impacts.

The rest of the paper is organized as follows. Section 2 presents the methods used in this study. Section 3 shows and discusses the results and indicators to measure scientific culture proposed. We conclude the paper in Section 4.

METHODS

Research questions that guide this paper are the following: i. What are information needs on scientific culture in Colombia?; ii. How can be measured scientific culture?; iii. What is the adequate structure for indicators of scientific culture?.

In order to answer these questions, a mix of methodologies is used (see figure 1). First, we review the literature on scientific culture and indicators related to this topic, that include programs and national initiatives to foster scientific culture in Colombia, and governance and policies related to this topic. Second, we made a series of interviews with staff members of Colciencias to determine requirements of measurement on scientific culture. Third, with this information, we built an information matrix to prioritise and categorize information and determine indicators with respective metrics that implies to determine a measurement variable (characteristics that can include different values) and unit of measure (number, average, percentage, rate, etc.), and sources according to relevance and cost-effectiveness of estimation. Fourth, from indicators formulated, these were prioritised and corroborated by Colciencias to establish structure and contents of system of indicators to validate with relevant experts and potential users of these information and indicators, which was made by internet through electronic form that includes objective of study, explanation of form, validation by dimensions and acknowledgements. Fifth, following to Godin (2000) and results of internal and external validation exercises an indicator system is proposed determining for every dimension of scientific culture indicators related to inputs, process, and outputs designed indicator fact sheets that includes definition, objective, sources aggregation levels, time series, and calculation methods for indicators proposed.

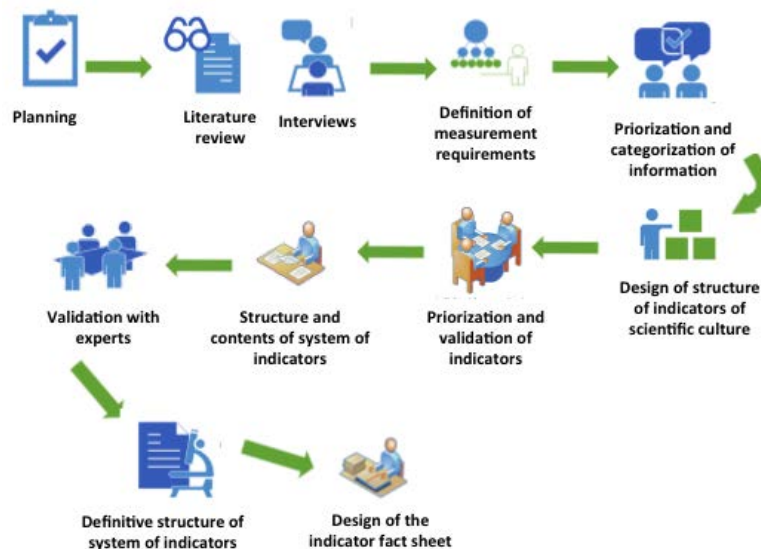


Figure 1. Methodological Route to develop and define a system of indicators to measure scientific culture in Colombia

RESULTS AND IMPLICATIONS

This study achieves formulate an indicator system from the definition of scientific culture and its dimension including two groups: i. Indicators related to general characteristics of scientific culture in Colombia denominated characterisation dimension of scientific culture, and ii. Specific indicators from three dimensions proposed according to national initiatives to strengthen scientific culture in the country (Training of human resource for science, technology and innovation, Public communication of science, technology and innovation, and Social appropriation of science, technology and innovation).

The indicator system is composed by 43 indicators through a multidimensional model that integrates different elements of scientific culture such as the individual and society establishing indicators to measure inputs, process and outputs in general form and specific initiatives for Colciencias and other institutions. Figure 2 shows the structure of indicators proposed.

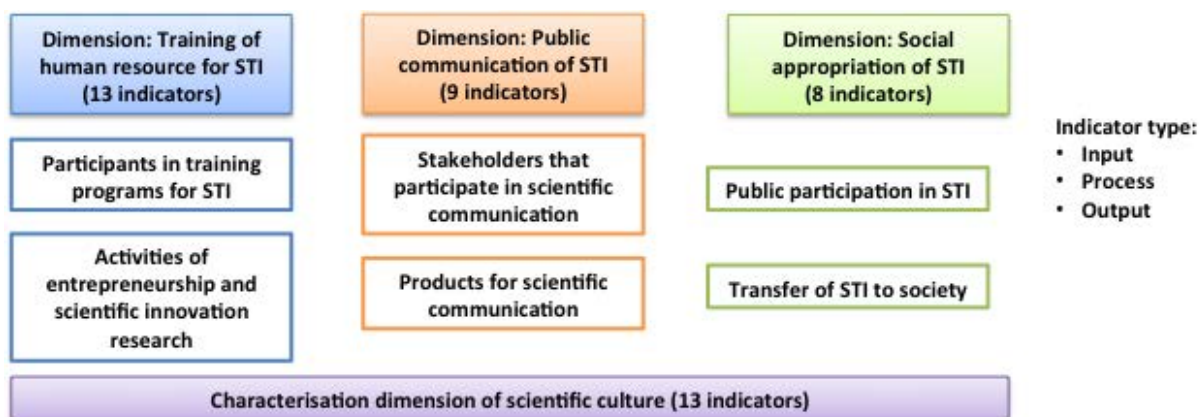


Figure 2. Structure of indicator system to measure scientific culture in Colombia

In characterisation dimension of scientific culture, it describes key elements on public policy related to this topic such as infrastructure for scientific culture, resources, projects and activities and results on population impacted with these initiatives (see table 1).

Table 1. Indicators proposed to measure characterisation dimension of scientific culture

Input indicators
Number of policy instruments that favour development of scientific culture.
Number of institutions that develop activities that encourage scientific culture depending on type of institution and geographic scope.
Number of science and technology fairs or spaces for dissemination of scientific culture depending on type of space and institution.
Number of researchers that work to foster scientific culture depending on area of knowledge and level of education.
Number of journalists and facilitators that work to foster scientific culture.
Number of networks that work to foster scientific culture.
Monetary investments in programs related to scientific culture.
Number of science centres established.

Number of science centres strengthened.
Process indicators
Number of projects that foster scientific culture depending on topic and the target population. Number of activities implemented to foster scientific culture depending on the target population and scope.
Output indicators
Number of products generated to foster scientific culture depending on type and the target population. Citizen participation by topics related to science, technology and innovation depending on type of activity and localization.

Indicators for dimension of training of human resource for STI allow to track to initiatives to promote this field in the country from basic education to the training of higher quality human capital (see table 2). These indicators include the following elements: i. Training of children and youths in scientific culture; ii. Programs of higher education to promote scientific initiation research; and iii. Process and developments of the new technology-based firms as a result of training in STI.

Table 2. Indicators proposed to measure Training of human resource for STI

Input indicators
Number of qualified teachers in STI in different levels of training. Number of financial support for formation process in master and doctorate depending on area of knowledge. Number of educational institutions that participate in programs that foster scientific culture.
Process indicators
Number of scientific research projects developed by children and youths. Number of teachers that participate in programs that foster scientific culture.
Output indicators
Percentage of individuals interested in science and technology careers. Percentage of university science and technology graduates. Number of students that participate in entrepreneurial activities of technology base depending on area of knowledge and level of education. Number of the new technology-based firms depending on type. Number of children and youths that participate in programs that foster scientific culture. Results of synthetic index of educational quality in educational institutions that participate in program that foster scientific culture. Number of undergraduate students that participate in initiatives of scientific initiation research. Number of products generated by undergraduate students that participate in initiatives of scientific initiation research.

Dimension of public communication of STI proposes indicators to measure different actions from diffusion of scientific research to public participation process where citizens help in the knowledge production and contribute in the public policy debate in topics on STI (Bubela, et al., 2009). These schemes imply multidimensional communication process among stakeholders where all have knowledge and decision-making capacity (see table 3).

Table 3. Indicators proposed to measure public communication of STI

Input indicators
Coverage (number of hours) of S&T programs on TV, radio, and in film. Number of calls to promote communication projects on STI.
Process indicators
Number of strategies that promotes the training of journalist or facilitators in process or dynamics of STI. Number of communication activities in STI that promote the interaction between scientists and society. Number of communication strategies in STI developed by different social groups.

Output indicators
Number of products of communication of knowledge developed by research groups depending on area of knowledge and type of product.
Number of individuals that consume products of communication related to STI.
Number of products of communication of knowledge directed to specific population.
Number of scientific communication contents that include to communities as co-authors.

Dimension of social appropriation of STI proposes indicators to measure public participation in topics on STI analysing level of participation, contribution to the discussion, evaluation and control of governmental actions related to these issues. Other important point, it is evaluate the process of transfer and sharing of knowledge that seek the solution of social problems where local knowledge, context, and the contributions of community are used (see table 4).

Table 4. Indicators proposed to measure social appropriation of STI

Input indicators
Number of instruments that promote the participation of researchers in process of social appropriation of STI.
Process indicators
Number of citizen participation instances in topics related to STI depending on the target population.
Number of projects of STI that generate activities of transfer and sharing of knowledge.
Number of projects of STI that response to social problems with community involvement depending on area of knowledge and the target population.
Number of projects of STI focused on population in vulnerable situation that promote transfer and sharing of knowledge depending on area of knowledge and the target population.
Number of research projects in topics on social appropriation of STI.
Output indicators
Number of researchers that have participated in the design of public policies related to STI depending on area of knowledge.
Number of individuals that intervene in citizen participation instances in topics related to STI depending on level of participation and population type.

These indicators are important to analysis trends and dynamic of scientific culture in Colombia from a monitoring that allow to determine if different public strategies are achieving a better positioning of science, technology and innovation in the society. Moreover, policy makers and decision-makers can use these indicators as a key input to generate effective science and technology policies from cost effective indicators.

CONCLUSIONS

This research proposes and develops an indicator system to measure scientific culture in Colombia from definition and dimensions of scientific culture established by Colciencias as governing body of science, technology and innovation policy, where it achieves to identify and determine information needs with their respective metrics.

The indicator system to measure scientific culture includes four dimensions: Characterisation of scientific culture (13 indicators), training of human resources for STI (13 indicators), public communication of STI (9 indicators) and social appropriation of STI (8 indicators). Therefore, this study proposes 43 indicators through a multidimensional model that include three types of indicators defined as input indicators (15 formulated) that measure resources (human and financial) dedicated to a particular program or intervention to promote scientific culture; process indicators (12 formulated) measure ways or methods in which program

services or strategies are provided to promote scientific culture in the country; and output indicators (16 formulated) measure the results, efficiency and effectiveness of different programs, sub-programs, agencies, and multi-unit/agency initiatives related to scientific culture in Colombia (Horsch, 1997), which allow an integral analysis and evaluation of dynamics and trends of scientific culture in Colombia.

These indicator system and model of measurement will allow to identify in a rigorous manner associated products to the promotion of scientific culture in Colombia that are made by different stakeholders and that in many occasions are not visible despite their valuable contribution to promote knowledge society, where can be defined these initiatives according to type, scope and effects in welfare. Findings of this study are important to establish a process of periodically measuring of the progress in empowering society with STI, define new strategies to promote the importance of STI in the society and lead to decision-making based on facts that strengthen a scientific culture in Colombia as a development engine.

Finally, it is important to measure periodically this indicator system with the aim that all stakeholders can establish trends and dynamics of scientific culture in Colombia and can contribute to design and formulation of new effective programs and instruments that promote knowledge-based society.

REFERENCES

- Departamento Administrativo de Ciencia, Tecnología e Innovación. Colciencias. (2010). *National strategy of social appropriation of science, technology and innovation*.
- Bauer, M., Allum, N., Miller, S., 2007. What can we learn from 25 years of PUS survey research? Liberating and expanding the agenda. *Public Understanding Science* 16, 79-95.
- Bubela, T., Nisbet, M., Borchelt, R., Brunger, F., Critchley, C., Einsiedel, E., Geller, G., Gupta, A., Hampel, J. 2009. Science Communication Reconsidered. *Nature Biotechnology*, 27, 14-18.
- Bucchi, M. 1998. *Science and the Media. Alternative Routes in Scientific Communication*, London and New York: Routledge.
- Bucchi, M. 2008. Of deficits, deviations and dialogues. Theories of public communication of science. Chapter 5. In *handbook of public communication of science and technology*. New York, Routledge.
- Colciencias, 2016. *National policy of Science, Technology and innovation, 2015-2025*. <http://www.colciencias.gov.co/noticias/consulta-el-borrador-del-conpes-de-pol-tica-de-ciencia-tecnolog-e-innovaci-n>
- Collins, H., Pinch, T. 1993. *The Golem: What Everyone Should Know about Science*. Cambridge: Cambridge University Press.
- Godin, B., Gingras, Y. 2000. What is scientific and technological culture and how is it measured? A multidimensional model. *Public Understanding of Science* 9, 43-58.

Horsch, K., 1997. Indicators: Definition and Use in a Results-Based Accountability System. <http://www.hfrp.org/publications-resources/browse-our-publications/indicators-definition-and-use-in-a-results-based-accountability-system>

OECD. 2015. *Science, Technology and Industry Scoreboard 2015 Innovation for growth and society*. http://dx.doi.org/10.1787/sti_scoreboard-2015-en



How user-innovators can be identified? Evidence collected from the analysis of practices¹

Konstantin Fursov^{*}, Alena Nefedova^{**} and Thomas Thurner^{**}

*ksfursov@hse.ru

Laboratory for Economics of Innovation, National Research University Higher School of Economics,
Myasnitskaya st., 9/11, Moscow, 101000, Russian Federation

** anefedova@hse.ru;*** tthurner@hse.ru

Laboratory for Economics of Innovation, National Research University Higher School of Economics,
Myasnitskaya st., 9/11, Moscow, 101000, Russian Federation

ABSTRACT

This paper studies how daily routines around media consumption, internet and technology-usage, product preferences or civic engagement mediate likelihood of being a user-innovator. Based on the differences in demographic characteristics of consumers and assimilation by them of certain daily routines we conclude that a deeper analysis of day-to-day activities can help distinguishing user-innovators from non-innovating peers. It is argued that innovation-related actions are rooted in learned behaviour, can be observed through the daily routines and tell us more on user-innovation experience. We suggest that no individual practice, but instead sets of practices taken in different economic, social and cultural environments can explain how innovations grow and disseminate through the entire economy.

INTRODUCTION

Early works on user-innovation asked how industrial products could emerge out of customer ideas (von Hippel, 1978). The importance of user-innovation has largely been argued through efficiency of product development (Hienerth et al., 2014) and benefits for national economies. Studies estimated the aggregate spending of user-innovators to be in the tens of billions of dollars annually (e.g., de Jong et al., 2015; Gambardella et al., 2015). Especially sports enthusiasts showed a very high willingness for spending time and money in their most favorite pass of time (Raasch et al., 2008; Hienerth et al., 2011).

A specific aspect of user-innovation studies paid great attention to the diffusion channels that user innovators choose to share with peers or to commercialize their findings (Ogawa and Piller, 2006; von Hippel et al., 2012). The share of user-innovators that diffuse their innovation has been estimated to be low, at around 12% (de Jong et al., 2015; von Hippel et al., 2011, 2012). This has been related to possible entrepreneurial opportunities the innovators intended to pursue. Others suggested that personality characteristics also have an influence on knowledge sharing (Matzler et al., 2008). Contrary to these findings, data out of Russia revealed a much higher rate of sharing (Fursov and Thurner, 2016). These findings were argued to be rooted in long-established practices in the day-to-day lives during the late Soviet Union, when goods supply in large parts of the country was at a sub-optimal level and user-innovation activities could play a role of a compensatory mechanism for non-market

¹ This study was conducted within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE) and supported within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program.

economic relations. This observed variation in sharing practices raises the question to which degree innovation-related actions are rooted in learned behavior more than in the psychological set-up of a person.

The concept of practice allows studying experiences of meaningfulness, as daily routines are the processes through which humans interact with the world around them. Hence, sociological theories have paid great attention to such practices considering them as an entire part of the “lifeworld” (Habermas, 1987), important socialisation mechanisms (Bourdieu, 1977, 1990), instruments of social control (Foucault, 1982), meanings that allow smooth performances of everyday life (Garfinkel, 1967), tools of production and reproduction of social order (Giddens, 1984). If the topic of practices and routines is in the focus of academic research, the question is mostly about how such practices can be alternated in order to be more environmentally sustainable or socially acceptable.

MOTIVATION OF RESEARCH AND RESEARCH QUESTION

In this paper, we study a group of variables derived from daily practices of media consumption, social networking, internet usage, civic engagement and some others to test their discriminatory power between Russian user-innovators and a group of non-innovating consumers. This comes from previous findings showing that information and skills for user-innovation are task-dependent (von Hippel et al., 2011; Jong et al., 2015; Lüthje et al., 2005), but user-innovators have been shown to be close followers of important market trends (von Hippel, 2005). Also they are sophisticated users of technologies and related products (Morrison et al., 2000; Luthje et al., 2005; Tietz et al., 2005). A specific interest rests on the use of Web 2.0 technologies through social networking sites, bulletin boards and online communities (Kietzmann et al., 2011; Ritzer and Jurgenson; 2010; Franke and Shah, 2003).

This paper follows the increasing interest in learning more about user-innovators and asks if practices and routines of user-innovators separate them from their non-innovating peers. As this study is based on a large data set derived from a public opinion survey in Russia, our results also feed back into the ongoing debate about the characteristics of user-innovators. Previous studies on the demographics of user-innovators have already revealed striking differences between user-innovators in western countries vs Russia. For example, data out of a Russian context suggest the presence of 9.6% of user-innovators, which far exceeds findings from other countries. Furthermore, Russian consumer-innovators are actively sharing their ideas. Almost 50% of the user innovators engage in such sharing activities. If the older cohort were taken out, the number would be even higher (Fursov and Thurner, 2016).

Russia is also an interesting case as its user-innovators act largely outside classical commercialization channels. Despite 20 years of reforms and attempts of modernization, Russia’s economy suffers from poor framework conditions such as low regulatory quality, questionable quality of institutions (Polischuk, 2013) or wrong incentives and stimuli resulting from flaws in Russia’s corporate governance models (Enikolopov and Stepanov, 2013). This puts the experience of Russia in stark contrast to other geographical areas where the focus rests greatly on entrepreneurship (e.g. Franke and von Hippel, 2003; Shah and Tripsas, 2007). This absence of easily accessible entrepreneurial routes makes the Russian experience even more interesting as they serve as a guideline for the many other countries in the world that find themselves in a similar situation.

This paper studies people in urban and rural community environments that modify or develop goods or services for their own benefit. Thereby, the study follows ideas developed by von Hippel (2005) and goes beyond conventional statistical frameworks, which require a connection to market-based activities. As the debate on whether the current definition is suitable to accommodate users that share knowledge with a peer group or community of

practice is ongoing, we believe that further insights also support including user-innovators (not only individuals) to the measurement framework (Gault, 2012).

METHODOLOGY

The data for this paper was derived from a large-scale public opinion survey in November 2014 within the framework of the Monitoring Survey of Innovative Behaviour of the Population (<http://www.hse.ru/en/monitoring/innpeople/>). The overall stratified sample consists of 1670 participants of 16 years and older, representative for Russian population by age, sex, education level, region (at federal district level), and city size. Data was gathered through face-to-face interviews. Selection bias for controlled social groups is not exceeding 0.03%. We targeted user-innovation on an individual level but not for „household sector innovators“ or unincorporated businesses (as suggested by Ferran, 2000).

The questionnaire covered the respondents“ experience in user innovation. Following von Hippel et al. (2011; 2012), we asked participants for a short description of their proclaimed user innovations (creation of new things or modification of existing products adopting them to respondents“ needs) in the last three years. We did not distinguish between the creation of new or the modification of existing products (unlike de Jong and von Hippel, 2009 or Pongtanalert and Ogawa, 2015). The questionnaire further captured, among other information, a list of daily practices, which we used as independent variables. We separated the respondents into user-innovators and non-user-innovators and applied a discriminatory analysis to study statistical differences between the groups.

FINDINGS

We first tested targeted media consumption and their predictability of user-innovators vs mere consumers. Mastering the English language is often an entry point to connect to a wider spectrum of topics and areas of interest outside the focus of Russian media coverage. Indeed, the data supports our assumption and shows a high and significant difference between user-innovators and others. Our data supports this view and shows that user-innovators are watching less Russian TV channels than non-users; however, we find it surprising that preference for foreign channels is not significantly discriminating between the groups. Interestingly from all other groups of media, user-innovators consume much more than non-innovators.

Our next variables were targeted the use of social networking sites (SNS). While the use of the most popular Russian SNS did not show any significant difference between user-innovators vs non-innovators, using an international SNS like Facebook did. The clarity of these findings is surprising. Firstly, Facebook provides most of its content in English, although the use of Russian is possible. Social network enthusiasts with limited language skills will probably revert to the Russian offerings. Those who do master English as a second language have also enjoyed a better education, which has been shown by previous studies.

The next five variables target Internet practices. User-innovators are much more acquainted with e-commerce practices and frequently buy goods online. Furthermore, user-innovators offer their own goods and services on more advanced platforms like internet-auctions. The greatest discrimination power was shown as the interaction with public administration like applying for a passport or other public services. User-innovators also more actively use the Internet as a communication tool for arranging services like appointments with a doctor. Online banking as a payment practice shows no significant difference between user-innovators and their non-innovating peers.

Furthermore, we asked if the use of technologies in daily practices do successfully discriminate between user-innovators and mere users. We chose for our study the practices

around transportation. For travelling with a car, we asked if our respondents use hardware like a GPS navigator. Using such devices indeed varied between the two groups. Interestingly though, using services like online maps or online information about traffic conditions have a lesser discriminatory power but are still significant.

All of the above might well point towards tech-enthusiasts that are greatly interested in the newest products and seeing what their new toys are actually able to. Hence, as a kind of control measure we ask the respondents if they seek to have new goods or services earlier than others do. Surprisingly, the enthusiasm for new goods and services showed no difference between the groups. There are, however, interesting differences in the choice of information that influence the buying decision. While non-innovating users prefer the advice of sales personnel, user-innovators look for shared experiences on Internet forums, product reviews or other independent experts' opinions. While both groups are equally paying attention to the price of a product or its brand, user-innovators are much more interested in whether a product they purchase is in fact environmentally friendly or is energy efficient in its use.

DISCUSSION AND CONCLUSIONS

Our research was especially motivated by remarkable differences in the willingness to share information and innovations between empirical findings from western countries and Russia. Based on these findings, we further delved into practices and routines as explanations for user-innovation.

The practices we included are largely connected with the use of certain technologies. One may argue that such practices would only hold true for user-innovators in technological areas, and especially software development would be an easily accessible field of user innovation. However, such skills surely are necessary to provide user-innovations in any technology-oriented field. On the other hand, using online information sources enables connecting with larger groups of likeminded enthusiasts. Previous findings from Russia revealed though a strong group of rural-based user-innovators that focus their attention on innovations around gardening and home decoration (Fursov & Thurner, 2016).

Previous papers have pointed out certain aspects of user innovators, like the higher education, their willingness to connect with like-minded individuals or their keen use of the latest gadgets. The present findings though are characterized by the high quality of our suggested model based on a set of practices in distinguishing user-innovators from non-innovators. Our paper shows that no individual practice, but instead a set of practices has a high likelihood to distinguish user-innovators. In our study, 73.5% of original grouped cases were correctly classified (for the results of the tests of equality of group means see Fursov et al., 2016). Given the high interest in identifying user-innovators, the practices we identified could serve as a promising starting ground for further investigation for specific "bundles (clusters) of practices" in different economic, social and cultural environments and how innovation growth could be supported through a wider dissemination of facilitating technologies.

For user-innovators, the most important basic requirement is access to materials and tools for innovation. Access to online shops helps a great deal to overcome limited availability of goods and services in rural territories. Hence, we stress the importance of available and affordable internet-connections for as many people as possible. Connecting to the global flow of ideas and actively exchanging information is vital for user-innovators who see internet-based technologies as a preferred means of communication. Ideas to ripe require a selected group of knowledgeable peers who voice concerns if there are any. These demands are especially important as good parts of developing countries especially in Africa still struggle with providing Internet access.

Another very promising research question could target the origins of these practices. Are these practices rooted in earlier socialization phases, like families or schooling, or have they been acquired at a later stage, e.g. through socialization processes at the workplace or amateur communities and memberships in clubs? A good deal of research on user-innovation has studied knowledge sharing in such amateur communities. These findings could be strengthened by deeper insights into the underlying practices of communities or sub-cultures respectively the larger cultural environment in which they are happening. However, as important as practices and routines are, individuals cannot always follow their routine ways. The question arises what are possible „breaking routines“, and how are these breaks interfering with innovative behaviour?

Some of the practices we have identified will only have the explanatory power in the Russian context (e.g. see the usage of international vs national social networking sites). Other variables might be less important in other countries. There could be practices that relate to certain cultural settings and spread over national boundaries, but loose significance elsewhere. To find out about practices of a national and regional importance, culturally bound practices and those with international significance, we rely on further studies on practices and routines.

REFERENCES

- Bourdieu, P. (1977). *Outline of a Theory of Practice*. Cambridge University Press, Cambridge.
- Bourdieu, P. (1990). *The logic of practice*. Stanford University Press.
- De Jong, J. P., & von Hippel, E. (2009). Transfers of user process innovations to process equipment producers: A study of Dutch high-tech firms. *Research Policy*, 38(7), 1181-1191.
- De Jong, J. P., von Hippel, E., Gault, F., Kuusisto, J., & Raasch, C. (2015). Market failure in the diffusion of consumer-developed innovations: Patterns in Finland. *Research Policy*, 44(10), 1856-1865.
- Enikolopov, R., & Stepanov, S. (2013). Corporate governance in Russia. *Journal of Institutional and Theoretical Economics*, 169(2), 355-381.
- Ferran L. (2000) The concept of the household sector in the 1993 SNA and further elaborations. *Household Accounting: Experience in Concepts and Compilation, Handbook of National Accounting*, Vol. 1, Chap. 1 (United Nations, New York), 19–55.
- Foucault, M. (1982). The subject and power. *Critical Inquiry*, 8(4), 777-795.
- Franke, N., & Shah, S. (2003). How communities support innovative activities: an exploration of assistance and sharing among end-users. *Research Policy*, 32(1), 157-178.
- Franke, N., & Von Hippel, E. (2003). Satisfying heterogeneous user needs via innovation toolkits: the case of Apache security software. *Research Policy*, 32(7), 1199-1215.
- Fursov, K. S. & Thurner, T. (2016), God Helps Those Who Help Themselves! A Study of User-Innovation in Russia. *Higher School of Economics Research Paper No. WP BRP 59/STI/2016*. Available at: <https://www.hse.ru/data/2016/05/16/1128809970/59STI2016.pdf>.
- Fursov K., Nefedova A., Thurner T. What user-innovators do that others don't: a study of daily practices *Higher School of Economics Research Paper No. WP BRP 62/STI/2016*. Available at: <https://www.hse.ru/data/2016/06/27/1116049975/62STI2016.pdf>.
- Garfinkel, H. (1967). *Studies in ethnomethodology*. Prentice-Hall, Englewood Cliffs, NJ.
- Gault, F. (2012). User innovation and the market. *Science and Public Policy*, 39(1), 118-128.
- Giddens, A. (1984). *The constitution of society: Outline of the theory of structuration*. University of California Press.
- Habermas, J. (1987). *Eine Art Schadensabwicklung Kleine Politische Schriften Vi*.

- Hienerth, C., Keinz, P., & Lettl, C. (2011). Exploring the nature and implementation process of user-centric business models. *Long Range Planning*, 44(5), 344-374.
- Hienerth, C., von Hippel, E., & Jensen, M. B. (2014). User community vs. producer innovation development efficiency: A first empirical study. *Research Policy*, 43(1), 190-201.
- Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, 54(3), 241-251.
- Lüthje, C., Herstatt, C., & Von Hippel, E. (2005). User-innovators and “local” information: The case of mountain biking. *Research Policy*, 34(6), 951-965.
- Matzler, K., Renzl, B., Müller, J., Herting, S., & Mooradian, T. A. (2008). Personality traits and knowledge sharing. *Journal of Economic Psychology*, 29(3), 301-313.
- Morrison, P. D., Roberts, J. H., & Von Hippel, E. (2000). Determinants of user innovation and innovation sharing in a local market. *Management Science*, 46(12), 1513-1527.
- Ogawa, S., & Piller, F. T. (2006). Reducing the risks of new product development. *MIT Sloan management review*, 47(2), 65.
- Polischuk L (2013). Institutional performance. In: Alexeev M., Weber S. (eds) *The Oxford handbook of the Russian economy*. OUP USA Oxford Handbooks in, Economics, pp 189–220.
- Pongtanalert, K., & Ogawa, S. (2015). Classifying user-innovators—An approach to utilize user-innovator asset. *Journal of Engineering and Technology Management*, 37, 32-39.
- Raasch, C., Herstatt, C., & Lock, P. (2008). The dynamics of user innovation: Drivers and impediments of innovation activities. *International Journal of Innovation Management*, 12(03), 377-398.
- Ritzer, G., & Jurgenson, N. (2010). Production, Consumption, Prosumption The nature of capitalism in the age of the digital „prosumer“. *Journal of consumer culture*, 10(1), 13-36.
- Shah, S. K., & Tripsas, M. (2007). The accidental entrepreneur: The emergent and collective process of user entrepreneurship. *Strategic Entrepreneurship Journal*, 1 (1-2), 123-140.
- Tietz, R., Morrison, P. D., Luthje, C., & Herstatt, C. (2005). The process of user-innovation: a case study in a consumer goods setting. *International Journal of Product Development*, 2(4), 321-338.
- Von Hippel, E. (1978). Successful Industrial Products from Customer Ideas. *Journal of Marketing*, 42(1), 39–49.
- Von Hippel, E. A. (2005). *Democratizing innovation*. Cambridge: The MIT Press.
- Von Hippel, E., De Jong, J. P., & Flowers, S. (2012). Comparing business and household sector innovation in consumer products: findings from a representative study in the United Kingdom. *Management Science*, 58(9), 1669-1681.
- Von Hippel, E., Ogawa, S., & De Jong, J. P. (2011). The age of the consumer-innovator. *MIT Sloan Management Review*, 53(1), 27.



Assessing youth engagement with a collaborative Index

Ramon Marrades^{*}, Filipa Pajevic^{**} and Simona Dobrescu

*ramon.marrades@uv.es

Cultural Economics and Tourism Research Unit. Faculty of Economics. University of Valencia. Avinguda Tarongers s/n 46022 València

** filipa.pajevic@mail.mcgill.ca. School of Urban Planning. McGill University. Macdonald-Harrington Building. 815 rue Sherbrooke Ouest. Montréal (Québec) Canada

As a response to the proliferation of student-led protests and movements across the globe, we, as part of an international platform for young planning professionals- Urbego-, have developed the Youth Engagement Index (YEI) that assesses the involvement of young generations (ages 18-34) in urban governance. Designed to include, and be improved upon by, a collaboration with relevant actors such as local municipal governments, academia, non-governmental youth organizations and the youth themselves, the YEI presents a unique opportunity to unveil weaknesses and opportunities for cities in terms of engaging their youth. Furthermore, the collaborative process highlights the value of having a recognized and engaged youth for future urban development and city life in general.

Through a qualitative and quantitative approach, the YEI looks beyond mediated youth-led protests by revealing more discrete, but not any less important, ways in which young generations could contribute to their cities. It is important to highlight and present these opportunities in a way that could be easily adopted and implemented through policy. Moreover, it is important to allow for better recognition and more acute inclusion of youth in urban governance.

The YEI evaluates engagement in a multi-dimensional way, addressing political, economic and social spheres of engagement. The series of participatory workshops conducted thus far have assessed 24 different variables of youth engagement in Medellin, London, Bucharest and Valencia. The variables include level of tolerance, safety, political trust and representation, and employment and funding opportunities. Together with a cohort of local government representatives, academia and youth, the variables were ranked, facilitating a discussion on state of policy and future implications. This paper presents the results of the YEI, along with considerations on the observed differences in perception and practices of urban governance in the aforementioned cities, and how they influence engagement in a broader sense. Finally, the paper argues that more collaborative processes such as the YEI are needed for a better understanding of youth involvement in an increasingly complex and interconnected global society.



Scientific Culture Measures: Challenges and New Perspectives

José Antonio López Cerezo*, Ana Muñoz van den Eynde**

*cerezo@uniovi.es

Professor of Logic and Philosophy of Science (Department of Philosophy), University of Oviedo, Teniente Alfonso Martínez s/n, Oviedo, 33011 (Spain)

**ana.munoz@ciemat.es

Researcher (Research Unit on Scientific Culture), CIEMAT, Av. Complutense, 40, Madrid, 28040 (Spain)

ABSTRACT

Since mid-twentieth century, efforts to promote scientific and technological development and engage the public in R&D process are increasing. Among those efforts, since the 1970s first in United States and then in United Kingdom and Europe, governments have funded surveys aimed at understanding the public attitudes toward science, scientists, and science policy. The *Science and Engineering Indicators* series of the National Science Foundation, or the European Community through its Special Eurobarometer on Europeans, science and technology, have shaped the research, measures and indicators of public understanding of science surveys. Examples are, at international level, surveys like *Scientific Culture in Ibero-American Countries* (2009, FECYT-OEI-RICYT), or the *International Study on Scientific Culture* (2012, BBVA Foundation); and at national level, surveys like *Social Perception of Science and Technology* (2002-2014) series, or the recent *Perception, Interest, Knowledge, and Actions (PIKA) Survey* (2014), both funded by Spanish Government through its Spanish Foundation for Science and Technology (FECYT).

During years, research on public understanding of science (PUS) has been shaped by standardized indicators and quantitative measures in order to enable international and longitudinal studies, and to ensure the representativeness of the population. Recent research on scientific culture suggests that the traditional approach, it is, PUS equals to high level of scientific literacy plus positive attitudes to science, is not enough, and a new approach is needed. Far away from PUS paradigm, scientific culture is taken as something that involves scientific knowledge, stablished one and controversial one; but also skeptical attitudes to science, and the willingness to use the scientific information in the decision making on issues regarding science and technology. Therefore, this panel on Scientific Culture Measures embraces new proposals oriented to improve a better comprehension of scientific culture, and ultimately, stimulate a more suitable research on public engagement with science.

The panel on Scientific Culture Measures is open to researchers and people concerned on reflections about scientific culture, its measures, the factors involved, public views of science and technology, etc. It may be of interest to other researchers who are concerned in exchanging proposals on new tools and resources to measure and empirically study social phenomena related to science and technology (innovation culture, risk culture, public R&D policies). Under STI Conference, this panel opens a space to share perspectives, encourage critical views and face new challenges in addressing issues of scientific culture. It also provides an excellent opportunity to introduce new research lines, especially in interdisciplinary contexts, among researchers from different fields and countries.

1. *Presentation / Introduction to the topic: What is scientific culture and what is not?*, José Antonio López Cerezo

2. *What does it mean to be scientifically literate?*, Belén Laspra

This contribution reviews the question "How much scientific literacy must have a citizen to meet the standards of being scientifically literate?" regarding to a broader concept, the scientific culture.

3. *New tools and indicators to measure scientific culture*, Ana Muñoz van den Eynde

Critical voices are emerging in PUS studies signaling to the need of a research reorientation due to methodological and theoretical deficits. PIKA Survey on Perception, Interest, Knowledge, and Actions related to science is a first contribution to this new approach.

4. *New cultural factors influencing the innovation measures*, María Cornejo Cañamares

Innovation can be influenced by certain cultural values, norms, attitudes and behaviors which may be considered as innovation culture. This contribution provides an interdisciplinary and critical view of innovation culture and its measure.

Agenda (Total length: 90 minutes)

5. *Presentation / Introduction to the topic: What is scientific culture and what is not?*, José Antonio López Cerezo (15 minutes).

6. *What does it mean to be scientifically literate?* Belén Laspra (15 minutes).

7. *New tools and indicators to measure scientific culture*, Ana Muñoz van den Eynde (15 minutes).

8. *Cultural factors influencing innovation measures*, María Cornejo Cañamares (15 minutes).

9. Open-ended discussion (30 minutes).

INFORMATION ABOUT THE ORGANIZERS AND PARTICIPANTS

ORGANIZERS

José Antonio López Cerezo is Professor of Logic and Philosophy of Science at the University of Oviedo and leader of the Social Studies of Science Research Group. His research focuses on STS, public participation, and more recently in the risk culture. He is the author and editor of numerous scientific publications on the field of STS and scientific culture research, and has been the main researcher of several national projects.

Ana Muñoz van den Eynde is researcher at the Research Unit on Scientific Culture at CIEMAT. She is Ph. D. and graduate in Psychology. Her research focuses on social perception, interest, knowledge, and behaviors about science, with special focus on environmental concern.

PARTICIPANTS

Belén Laspra is researcher at the Social Studies of Science Research Group of the University of Oviedo. She is Ph. D., and her research focuses mainly in the fields of scientific culture and public understanding of science.

María Cornejo Cañamares is researcher at the Research Unit on Scientific Culture at CIEMAT. She is Ph. D. in Social Studies in S&T. Her research interests focus on innovation and sustainability.

Preferred number of participants: 20

Special requests/equipment needs: computer and projector

REFERENCES

Cámara Hurtado, M. & López Cerezo, J. A. (2010). Political Dimensions of Scientific Culture: Highlights from the Ibero-American Survey on the Social Perception on Science and Scientific Culture. *Public Understanding of Science*, 21, 369-384.

Cámara Hurtado, M. & López Cerezo, J. A. (2015). “La población española ante el riesgo y las aplicaciones de la ciencia. El caso de los procientíficos moderados. FECYT, *Percepción social de la ciencia y la tecnología 2014*. (pp. 165-188).Madrid: FECYT.

Muñoz van den Eynde, A. (2015) “Factores que contribuyen a construir la imagen pública de la ciencia. La relación entre percepción, interés y conocimiento”. FECYT, *Percepción social de la ciencia y la tecnología 2014*. (pp. 15-44). Madrid: FECYT.

Muñoz van den Eynde, A., Cornejo Cañamares, M., Díaz García, I. & Muñoz, E. (2015). Measuring innovation culture: Development and validation of a multidimensional questionnaire. *Advances in Research*, 4 (2), 122-141.

Muñoz van den Eynde, A. & Lopera Pareja, E. (Coords.) (2014). *La Percepción Social de la Ciencia. Claves para la Cultura Científica*. Madrid: Los Libros de la Catarata.

Muñoz, A., Moreno, C. & Luján, J. L. (2012). Who is willing to pay for science? On the relationship between public perception of science and the attitude to public funding of science. *Public Understanding of Science*, 21 (2), 242-253.

Laspra, B. & Muñoz, E. (Coords.) (2014). *Culturas científicas e innovadoras. Progreso social*. Buenos Aires: Eudeba.

CHAPTER 13

Indicators in Regional and National Contexts



Indicators of the Knowledge based Society: Comparison between European and Latin American countries.

Daniel Villavicencio*, Alberto Morales** Marcela Amaro***

* dvillavic@correo.xoc.uam.mx, Univ. Autónoma Metropolitana-Xochimilco, Calzada Del Hueso, #1100, Col. Villa Quietud, Mexico City C.P.04960

** Almorlaesanchez@gmial.com, Fac. de Economía, UNAM, Circuito interior s/n, Ciudad Universitaria, Coyoacán, Mexico City C.P. 04510

*** marxelaser@hotmail.com Univ. Autónoma Metropolitana-Xochimilco, Calz. Del Hueso, #1100, Col. Villa Quietud, Mexico City C.P.04960

ABSTRACT

There has been a great deal of attention paid to measuring Information Society developments. There have been efforts to develop new statistics and systems of indicators to measure the diffusion of new information technologies in business and to examine levels of use and styles of use (e.g. e-Commerce). These efforts are ongoing and provide valuable material with which to compare different countries, regions and industrial sectors.

Other features of the Knowledge based Society have also attracted a great deal of attention. Indicators are an important instrument for monitoring the dynamics of the Knowledge based Society and to generate information for better policy interventions. In this sense, the main purpose of our research has been to understand the differences between European and Latin American countries in accomplishing their strategies and policies to access to what is called the Knowledge based Society. The central focus of our work was the assessment of the need for an improvement of Knowledge Society indicators and for statistical capacity building in order to capture the dynamics of the Knowledge Society in LAC countries and to make them comparable to EU and OECD standards. For doing so, we have chosen three different approaches.

The approach was to analyze the use of indicators, mainly based on the OCDE and World Bank methodologies, which are supposed to show the building of knowledge capacity of societies. The first section provides an overview on existing Knowledge Society indicators by carrying out a review of the current status of Knowledge Society statistics in some LAC countries. Besides the differences shown between the European countries and the LAC ones, we also identify the gaps in the construction of indicators amongst the Latin American countries.

INTRODUCTION

The blistering pace of technological progress has huge repercussions on society. Breakthroughs in microelectronics, cybernetics and telecom that increase data transmission speed and capacity are occurring simultaneously with substantial cost reductions, which in turn are leading to more widespread use of the new technologies. These changes in the global economy have given rise to the term *Knowledge-based Economy* (OECD, 1998), which underscores the use of knowledge as a factor of economic growth.

Different kinds and dimensions of knowledge is considered decisive issues in the economic growth of the advanced economies. There have been efforts to build new statistics and indicators to measure its progress in society. These efforts provide valuable material to compare different countries, regions or even industrial sectors. Other features of what we could call the Knowledge based Society (KbS) have attracted a great deal of attention, and some international organizations have propose methodologies to measure them.

Indicators are important for monitoring the dynamics of processes and to generate information for better policy design. In this sense, the main purpose of our research has been to understand the differences between European and Latin American countries (LAC) in accomplishing their strategies and policies to access to what is called the Knowledge based Society (KbS).¹ The central focus of our work was the assessment of the need for an improvement of KbS indicators and for statistical capacity building in LAC countries, as well as the comparison whit those used by European Union. In this sense, our analysis is based only on the statistical data available on government agencies and other institutional sources.

1.- The Knowledge-based Society

The term “knowledge society” refers to a broad spectrum of conditions and features that characterize a society, such as infrastructure, industries, tools, knowledge and research projects relating to new technologies and their widespread use by society. Access to telecommunications, information, as well as to the development of new processes and products, are other features of KbS. Production, manufacture and the way knowledge is used are all crucial factors as they define how social relations and activities are structured. In more general terms, the KbS involves cultural change by which society is immersed into a process of profound transformation implying not only the use of technologies but also a change of mentality and behavior based on the use of knowledge in the social, political and economic fields that provide benefits for society as a whole.

Sörlin and Vessuri (2007) describe the term “knowledge society” as an aspiration, more than a specific fact. It is a hypothetical society in which knowledge becomes society’s primary input for in economic growth and for social development. The terms “knowledge society” and “information society” are sometimes confused, and it is therefore necessary to differentiate them. Information is not the same as knowledge; the latter includes cognitive categories, codes for interpreting information, as well as tacit and heuristic skills for research.

The “knowledge society” refers primarily to a type of collective knowledge that has been encrypted in several aspects of the social system and, becomes visible or usable for solving specific problems. The knowledge society comprises different components, such as: the

¹ Here we present our reflection in the frame of the Network Convergence of Knowledge for Society based on the key findings of of *Eulaks* project (Connecting Socio-Economic Research on the Dynamics of the Knowledge Society in the European Union and Latin American and Caribbean Countries) FP7 grant agreement N° 217190.

education system, the socio-institutional system, public policy on science, technology and innovation, infrastructure, specialized human resources and the cultural basis that makes it possible to set joint goals for society involving the use of knowledge.

2. The necessity to measure the knowledge-based society

Information generated by scientific, technological and innovation-related activities can be used to monitor, prospect, research or determine the course of public policy in Science and Technology. The aim of this is to assess the progress made or the gaps existing between different societies around the world by considering that the KbS is an advanced state in which the evolution of the population and its economic and institutional structures has improved the standard of living. For this purpose we need to define variables to measure different positions and processes and their evolution.

The range of available indicators covers activity, performance, financing, invention, innovation and dissemination of knowledge, technologies, practices, infrastructure and the development of human resources. A suitable measuring system provides the tools required to identify areas that need to be developed as well as the strengths of a given country, region or sector.

However, the measuring of knowledge is not free of problems. One such problem is that the different components of knowledge are heterogeneous and there is no way of comparing them. Another problem is that a substantial proportion of knowledge is not observable (tacit knowledge) and it becomes difficult to register it. Thirdly, there is no model to measure the relationship between knowledge generation (*input*), and its economic effects (*output*). Lastly, measuring knowledge stocks is practically impossible. One thing that can be measured is the expenditures on knowledge generation activities, principally R&D and specific outputs like patents, publications and new products.

In order to gauge the progress made by different countries towards knowledge, it is crucial to have some quantification criterion that provides a means for comparison and for assessing results. This is still a difficult task because of the issues mentioned above, nonetheless some major breakthroughs have been made in this regard.

3. Some indicators of the Knowledge-based Society

Here we briefly present different methodologies and three groups of indicators as regards the measurement of the KbS. Our aim is to analyse the dimensions involved, their variables and assets.

World Bank Indicators.- The Bank's *Knowledge for Development* (K4D) program issued a global report of 140 countries with regard to the progress made in devising knowledge-based development strategies. As a result of globalization and the recent technological revolution, countries wishing to become part of the knowledge society need to strengthen their educational base, innovation system, communications and information infrastructure at the same time.

The indicators used have been built on four pillars defining the prerequisites for joining the knowledge society. The first pillar is the economic and institutional system, whose primary role is to provide incentives for the generation and acquisition of new knowledge, and for using it in the economy. The second pillar is education and skills necessary for individuals to be able to generate successful cognitive processes. Thirdly, there is the infrastructure of information and communications, which makes it possible to disseminate information throughout society as a whole. Lastly, the innovation system needs to be strengthened in order to allow the knowledge stock to be used and provide solutions for local issues.

Countries' knowledge performance is assessed by measuring 83 variables related to the four abovementioned pillars. This methodology can be simplified by measuring the 14 most important performance variables only, and its advantage lies in the international comparability of a series of indicators in a number of performance categories. The 83 variables that feature in this index can be divided into six categories: economic performance, economic regime, government, innovation system, education, gender and the Information and Communication Technologies sector. The idea is to be able to make discriminate selections from the 83 variables in order to obtain a series of substantive indicators that will provide a basis for comparison with different countries in Latin America with regard to their progress towards the knowledge society, in accordance with patterns defined by the European Union. The proposed selection is as follows.

Indicators proposed by the European Union.- In order to have a basis for comparison with Latin American countries, we propose to turn to the indicator matrix formulated by *The European Foundation for the Improvement of Living and Working Conditions*, which provides a good idea of the progress made and the requirements for creating societies of knowledge.

First of all, we can find indicators of the Statistical Office of the European Union (Eurostat), which generates and standardizes statistical data on member states. Then there are the data of the UNESCO Institute for Statistics, which provides data on communications users. There are also the data of the SIBIS project (*Statistical Indicators Benchmarking the Information Society*) that is part of the EU Information Society Program, which generates indicators on knowledge-intensive activities. In addition, there are OECD-proposed indicators including data on the use of computers and employment in member states. Data from the European Continuing Vocational Training Survey (CVTS) are also incorporated. And, lastly, there are the International Law Committee (ILC) and the International Labor Organization (ILO), both providing data on the GINI index and work-related productivity. Indicators are divided up between prerequisites for establishing a knowledge society, and quantifiable results derived from these variables.

The prerequisites variables are:

- a) Infrastructure and resources
- b) Socio-economic prerequisites
- c) Policies (government involvement).

The results variables are:

- a) Innovation capabilities
- b) Labor flexibility.
- c) Online applications
- d) Data on welfare and satisfaction.

The table below gives a breakdown of the main indicators measured by the European Union in connection with the development of KbS. Some of the variables put forward by the OECD for measuring the knowledge economy are taken into account.

Table 1, Knowledge Society Indicators

Prerequisites for the Knowledge Society		Indicator	Source
Infrastructure and resources	Media	Mobile phone subscribers	Eurostat
		Internet users	Eurostat
		Internet hosts	Eurostat
		Personal Computers	Eurostat
		Daily newspapers	UNESCO
		Television receivers	UNESCO
		Radio receivers	UNESCO
		Email users	SIBIS
		Email use and networking	SIBIS
		Broadband Internet access	OECD
		Internet access drop-outs	SIBIS
	Education	Pupil / teacher ratio	UNESCO
		Number of teaching hours per year in public institutions, by level of education	OECD
		Everyday computer availability at home	OECD
		Everyday computer availability at school	OECD
		Computers connected to the Internet at school	OECD
Employment	Standardised overall unemployment rate	Eurostat	
	Unemployment rates by level of educational attainment of 25 to 64-year-olds (lower and/or upper secondary and post-secondary non-tertiary education)	OECD	

Socio-economics		Unemployment rates by level of educational attainment of 25 to 64-year-olds	OECD
	Training and skills	Employees' participation in company-provided courses	CVTS
		Training enterprises	CVTS
		Enterprises evaluating the effect of CVT courses	CVTS
		Employees practising e-learning	SIBIS
		COQS-index of perceived digital literacy	SIBIS
		GINI-index on income disparity	ILC
	Social inclusion	DIDIX: Digital divide index	Empirica
		Effects of security concerns on e-commerce	SIBIS
		Governmental expenditure on R&D	OECD
Outcome variables		Indicator	Source
Applications, outputs, markets	Innovation ability	Patent applications	Eurostat
		Expenditure on R&D	Eurostat
		Labour productivity	ILO
		Employees in third sector	Eurostat
		Adaptability of work arrangements index	Empirica
	Work flexibility	Spread of tele-work (all types)	SIBIS
		Workers practising tele-cooperation	SIBIS
		E-commerce use	SIBIS
		E-health:	SIBIS
			SIBIS
		Gross domestic product (GDP)	Eurostat
	Wealth and satisfaction	Perceived job satisfaction	SIBIS
		Perceived job security	SIBIS

Standardized indicators for Latin America.- In addition to the indicators proposed by international organizations like the World Bank and UNESCO, there are a few sources of information whose role is to standardize data obtained from the ministries of statistics in Latin America. The most important one is the Network of Science and Technology Indicators (RICYT), a network spanning Latin America, Spain and Portugal whose main purpose is the creation and standardization of indicators.

RICYT encompasses the national science and technology agencies and Ministries in every Latin American country plus the ones in Spain and Portugal. The indicator comparison matrix consists of four general categories: context indicators, input indicators, product indicators and innovation indicators, divided up into different subcategories as shown in the table below. Nevertheless, if we bear in mind that the aim here is to measure the progress of the knowledge society in different countries, then these indicators are insufficient by themselves.

Table 2, RICYT: Standardized Indicators for Latin America

Context Indicators	Population	
	Labour Force	
	GDP	
Input Indicators	Financial Resources	Expenditure on Science and Technology
	Human Resources	Personnel on S&T
		Number of Researches
		Number of Graduates
Output Indicators	Patents	Patent applications
		Patent grant
		Dependence rate
		Self-Sufficiency Rate
		Invention Coefficient
	Bibliometric Indicators	Registered Publications
		Publications per inhabitant
Selected Innovation Indicators (Only for selected countries based on a survey)	Enterprises with R&D Departments	
	Enterprises with Innovation Activities	
	Enterprises with R&D Activities	
	Enterprises with Output Improvement	
	Enterprises with Process Improvement	
	Enterprises to turn to a consultancy	

Some of them, basically the ones referring to financial and human resources for R&D, patents and publications, can be used as result variables. One thing the RICYT does not measure is the requirements for the KbS. One noteworthy drive conducted by the RICYT in this regard is the measuring of select innovation indicators, even though there are some major limitations: they are only available for the region's bigger countries, and they are obtained through surveys that whose results cannot be easily generalized.

In order to be in synch with the construction of the knowledge society indicators proposed by the EU, the RICYT may standardize data on the training and cognitive skills of workers (e-learning, digital literacy, etc.), as well as their social inclusion (disparity of income). Furthermore, as

result indicators, it is also necessary to define variables relating to electronic applications (e-government, e-commerce) and new labor market conditions (remote work). While the RICYT generally focuses on standardizing science and technology indicators, the capacities acquired in building indicators allow it to expand its scopes for the purposes of creating knowledge society indicators. Yet, the main effort to be done involves the different Ministries, agencies and public institutions in LAC as they might produce new statistical information.

5.- Comparison of indicators from EU and LA countries.

All indicators available for the countries analyzed in our research are presented in a broad table, in which European indicators are used as a basis for comparison (Villavicencio et al, 2012). We had difficulties in following the periods when information is collected in some countries. This becomes a problem given that only international agencies and organizations have data for common periods. Furthermore, even if comparisons can be made between Mexico, Brazil and Uruguay, the years can vary for countries that do have the required information.

Incompatibility among indicators and the periods studied makes comparison difficult and poses new challenges, even though efforts are being made to measure the progress or shortfalls in the achievement of a KbS, it is often not possible to quantify the accomplishments of each LA country. The absence of parameters for identifying advances or setbacks means that the entire wealth of information compiled by national and international organizations is actually of little use.

The first thing that stands out in our comparison exercise is the particular emphasis Latin American countries place on indicators relating to results in the field of innovation (more disaggregated), such as the number of patents and spending on R&D. Nonetheless, countries in LA do not consider performance at work and staff training variables as an important factor in measuring the progress of the KbS. In general terms, requirement variables for the KbS are rarely included in the LA measurements; and where they are included, the information is highly disperse.

Our analysis revealed a huge disparity between LA countries with regard to information systems. While Brazil and Mexico have acceptable and functional systems and make valuable information available, in other countries such as Venezuela, information systems are still relatively rudimentary in its content, as well as limited in access and functionality. The disperse nature of information platforms and the lack of any shared conceptual basis among the LA countries make it difficult to perform comparisons.

6.- Conclusion

The analytical formulation of variables that can be used to measure the access to the *knowledge based society* is still a topic pending international discussion. There is a gap between the EU and LA regarding the definition of variables and indicators. The EU has progressed, managing a set of specific variables and standardize the available data of member states, which is not the case of LA. What is being measured in LA refers to knowledge result variables, and though significant

headway has been made in standardization through RICYT, there are still no signs of a determined drive to make analytical and conceptual progress towards setting suitable parameters for the knowledge society concept.

The Latina American countries analyzed have developed indicators that are often not compatible for the purposes of comparison and more important is that they cannot show the advancement of the society in acquiring and using knowledge for different purposes. An analysis of the path followed by European countries in the generation of variables may be a very useful exercise for LA, not just for the purpose of compiling the information available but also for defining the course to be followed by future measurements. This might also help governments to foresee new or better policies that can enhance learning capabilities of their societies in academic institutions, enterprises or public agencies, as well as individuals in general.

REFERENCES

- CIA World Factbook <https://www.cia.gov/library/publications/the-world-factbook/>
- European Foundation for the Improvement of Living and Working Conditions (2004) Advancement of the Knowledge society comparing Europe US and Japan..
- European Foundation for the Improvement of Living and Working Conditions (2003) Handbook of Knowledge Society Foresight..
- Observatorio para la Sociedad de la Información en América Latina y el Caribe (OSILAC) www.eclac.org/SocInfo/OSILAC/
- OCDE (1996) “The Knowledge Based Economy”, Paris
- RICYT: <http://www.ricyt.org/>
- SIBIS (2003) New eEurope Indicator Handbook.
- Sörlin, S. y Vessuri, H. (2007). Knowledge Society vs Knowledge Economy: Knowledge, Power, and Politics. New York: Palgrave MacMillan.
- Statistical Indicators Benchmarking the Information Society <http://www.sibis-eu.org/>
- UNESCO (2014) Institute for Statistics <http://www.uis.unesco.org/>
- World Bank, (2012), Knowledge Assessment Methodology, en <http://www.worldbank.org>
- World Bank (2014) “Knowledge for Development program”, en <http://www.worldbank.org/>
- VILLAVICENCIO D., A. Morales y M. Amaro, (2012), “Indicadores y asimetrías sobre sociedad basada en el conocimiento en A. Latina”, *Rev. Perfiles Latinoamericanos*, Flacso, México, vol. 20, pp. 63-96.



Measuring cross-border regional STI integration ¹

Teemu Makkonen*

*t.makkonen@surrey.ac.uk

School of Hospitality and Tourism Management, University of Surrey, Rik Medlik Building, Guildford, GU2 7XH (United Kingdom)

ABSTRACT

Earlier quantitative studies on cross-border regional integration processes have commonly neglected science, technology and innovation (STI) indicators: even the most notable example of a composite indicator approach to measuring cross-border regional integration, i.e. the Oresund index, lacks a sub-category for STI. Consequently, by ignoring cross-border innovation and knowledge flows, the Oresund integration index fails to take into account one of the most important drivers of economic growth in cross-border regions. Therefore, a new composite STI indicator (sub-category) was introduced to strengthen the Oresund integration index. This was compiled from patent, publication and collaborative R&D project data. The findings show that this index performs reasonably well in depicting STI integration, while at the same time remaining simple and straightforward enough to be adopted in other cross-border regions.

INTRODUCTION

Cross-border regions (CBRs) and their integration processes have been in the eye of political and scholarly debates for decades. However, measurement problems, related to data availability issues (not the lack of data per se, but the laborious process of gathering this data) and the choice of appropriate indicators to describe integration, have persistently hampered the quantitative investigation of these regions. Therefore, most quantitative studies on the subject have been restricted to studying a distinct feature of integration, such as labour markets (Schmidt, 2005), with a limited set of individual indicators. An interesting exception has been developed in the Danish-Swedish CBR of Oresund (Öresund for Swedes and Øresund for Danes); namely the Oresund integration index (Öresundskomiteen, 2013) – henceforward referred to as “the Index”.

The local authorities (that is, the Oresund Committee) responsible for raising awareness as well as studying and facilitating cross-border regional integration have collected data on various sub-categories of cross-border regional integration dating back to 2000, when the Oresund bridge (and a tunnel) – henceforward “the Bridge” – crossing the Oresund strait – henceforward “the Strait” – was opened, which significantly reduced travel times across the Strait, compared to ferry-traffic. In addition to the Oresund region being one of the most commonly used examples of cross-border regional integration (Nauwelaers, Maguire & Ajmone Marsan, 2013), the Index is exceptional since it is the only example of a CBR, where a time-series approach has been employed to study cross-border regional integration, together with composite indicators. In contrast, the other existing examples of studies focusing on cross-border regional integration commonly apply a cross-sectional approach with fixed years

¹ This work was supported by the Marie Curie Actions (Intra-European Fellowship for career development) within the Seventh Framework Programme (FP7) of the EU under Grant PIEF-GA-2013-624930

of analysis and a limited number of indicators (e.g. BAK Basel Economics, 2008; Decoville, Durand, Sohn & Walther, 2013). The Index uses a weighting scheme that allows the inclusion of various indicators of integration into a single composite indicator. This Index is then compared to the base year, 2000 (i.e. index year 2000 = 100), to indicate, whether the region has moved towards, what can be labelled, a more integrated CBR or has drifted apart despite the improvements in infrastructure and the political will laid on promoting integration. The Index consists of five distinct sub-categories of integration, containing from three up to five individual indicators, including: 1) labour markets, 2) housing markets, 3) business, 4) culture and 5) transport and communication sub-categories. The five sub-categories have equal weights in the total Index. Each of these sub-categories has between three to five individual indicators. The statistics applied to compile the basic indices are mainly derived from the Öresund database (Örestat, 2015). Additionally, to remove spurious trends and cyclical movements, the indices are adjusted by comparable indices that reflect the overall domestic developments in Denmark/Sweden. In short, the basic index is divided by the comparable index to obtain the adjusted basic index (OECD, 2013; Öresundskomiteen, 2013).

Despite being a notable example of time-series data and composite indicator approaches to cross-border regional integration, the Index still lacks a component (or sub-category) that takes into account what is arguably one of the most important drivers of regional economic development and competitiveness, namely the component (or sub-category) of STI. In short, as noted by Nauwelaers et al. (2013), the Index fails to capture cross-border knowledge and innovation flows, against a background of STI in cross-border regional integration and economic development in CBRs having been highlighted as essential in, for example the emerging field of cross-border regional innovation systems (CBRIS) literature (Lundquist & Tripl, 2009).

Therefore, the aim of this paper is to strengthen the Index by adding a new sub-category, which specifically addresses the essential parts of the economy related to STI indicators in a cross-border context. The new sub-category was constructed from relevant indicators, also recommended by the OECD (2013), collected from existing databases, including: 1) cross-border co-patents [OECD's Regional Patent (REGPAT) database], 2) cross-border co-publications [Web of Science (WoS) database] and 3) cross-border research and development (R&D) projects [Community Research and Development Information Service (CORDIS) database]. In line with this, an appropriate weighting scheme was suggested. Data was gathered from the year 2000 onwards corresponding to the Index. Subsequently, the composition and weighting scheme of the Index and the new STI sub-category are discussed, followed by the findings from the composite STI index. The results are presented together with comparisons to existing data on cross-border regional integration in the Oresund region. Thus, the paper aims at improving the Index and function as a reference for other regions and researchers interested in further quantitative works on CBRIS to apply, improve and develop composite indicators for measuring cross-border regional integration. This directly corresponds to the research agenda set by Nauwelaers et al. (2013) who stated that: "a more innovation-driven Oresund would need to be supported by an extension of the coverage of the Öresund database and a deepening of Örestat's work to cover innovation" (p. 10).

THE COMPOSITE STI INDEX

Cross-border co-patents measure the level of technological cross-border collaboration among firms, R&D centres and organisations in a given CBR (OECD, 2013). Cross-border co-patents have been rarely used when discussing the cross-border regional integration process of

the Oresund region (Moodysson & Jonsson, 2007). This is probably largely due to data availability issues related to the geographical scope of the Oresund region, which does not fit well with comparison made on the “official” NUTS-regions level: the Danish side has two NUTS-2 regions comprised of six NUTS-3 regions whereas the Swedish side is covered by a single NUTS-3 region. Therefore, this kind of data is not readily available from, for example, Eurostat’s statistical databases. Here the data for basic and comparable indices were gathered by using the REGPAT database by searching the address details in the inventor field of patents applications to identify first, cross-border co-patents between the Danish and Swedish sides of the CBR and second, the total number of cross-border co-patents between Denmark and Sweden (Table 1).

Cross-border co-publications measure the level of scientific cross-border collaboration among research institutions in a given CBR (OECD, 2013). The indicator has been in a wide use, largely due to the existence of convenient publication databases, when measuring cross-border regional integration processes in certain industries and scientific fields in the Oresund region (Coenen, Moodysson & Asheim, 2004; Hansen & Hansen, 2006; Moodysson & Jonsson, 2007; Hansen, 2013). Here the data for basic and comparable indices were collected from the WoS database. The names of the municipalities and towns belonging to the Oresund region were used to identify first, cross-border co-publications between the Danish and the Swedish side of the CBR and second, the total number of scientific article publications in the CBR (Table 1).

Cross-border collaborative R&D projects measure the intensity of cross-border collaboration among research organisations in a given CBR (OECD, 2013). To the best of the author’s knowledge this type of data has not been utilized previously in analyses of cross-border regional integration in the Oresund region. As with patents, this is probably principally due to data availability issues. However, the data can be mined from existing sources. Here the primary public repository comprising information on European Union (EU) funded R&D projects, i.e. the CORDIS database, was utilized to produce the basic and comparable indices. The geographical locations and names of the participating organisations were used to identify first, the number of (starting) cross-border collaborative R&D-projects between the Danish and Swedish side of the CBR and second, the total number of (starting) R&D projects in the CBR (Table 1).

As in the case of the Index, the individual indicators in the STI sub-category were assigned weights to reflect their importance *vis-à-vis* innovation and (potential) further regional economic development. Based on existing knowledge of the “pay-offs” from innovation related activities, the weighting scheme takes into account that while patents are not necessarily linked to actual innovations, introduced into the market, they have been commonly utilized to depict the output side of innovation in economic analyses on the geography of innovation. In contrast, R&D and scientific publications have usually been seen as input indicators of innovation (Hagerdoorn & Clodt, 2003; Makkonen & van der Have, 2013; Carvalho, Carvalho, & Nunes, 2015). Therefore, while acknowledging the simplicity of the solution, the individual indicators of the new STI sub-category receive the following weights: co-patents; 50%, co-publications; 25% and collaborative R&D projects; 25% (Table 1). The weighting scheme applied here is generally in line with the existing literature: for example, with the weighting scheme, based on fuzzy set theory and survey results of expert opinions in the field of STI, proposed by Moon and Lee (2005).

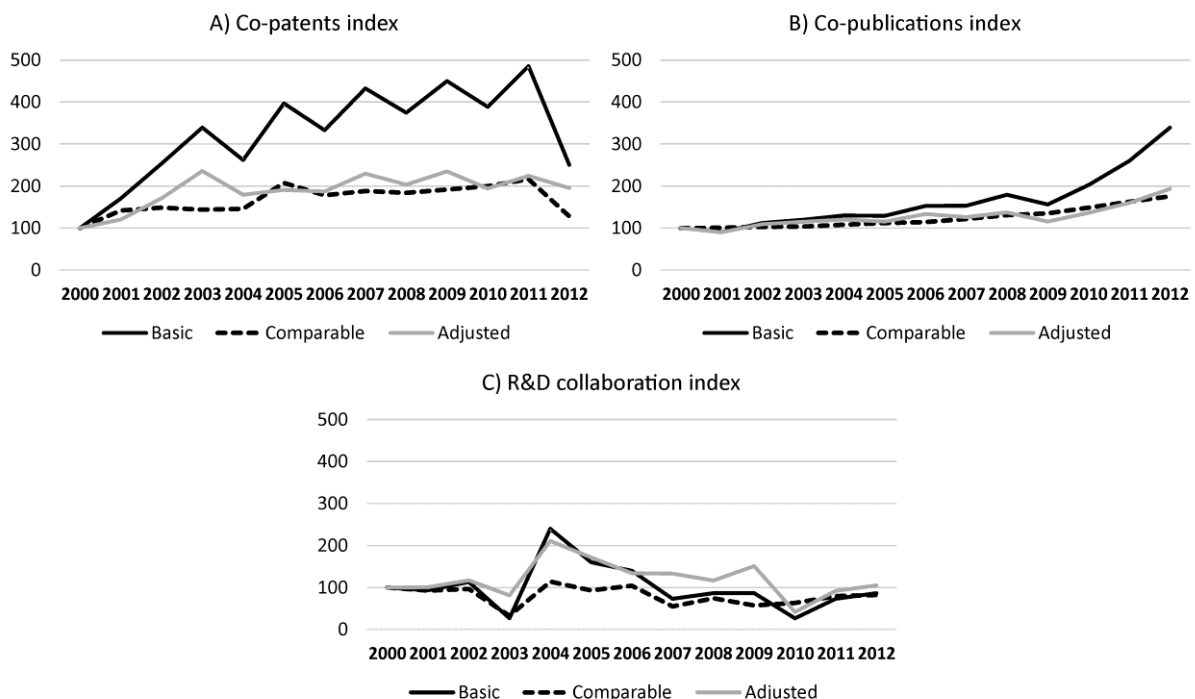
Table 1. Basic and comparable indices of the STI sub-category.

Weights	Basic index	Comparable index
50%	Cross-border co-patents across the Oresund strait	Cross-border co-patents between Denmark and Sweden
25%	Cross-border co-publications (scientific articles) across the Oresund strait	Total number of scientific article publications in the Oresund region
25%	Cross-border collaborative R&D projects across the Oresund strait	Total participation in R&D projects in the Oresund region

The composite STI index with comparisons to the Oresund integration index

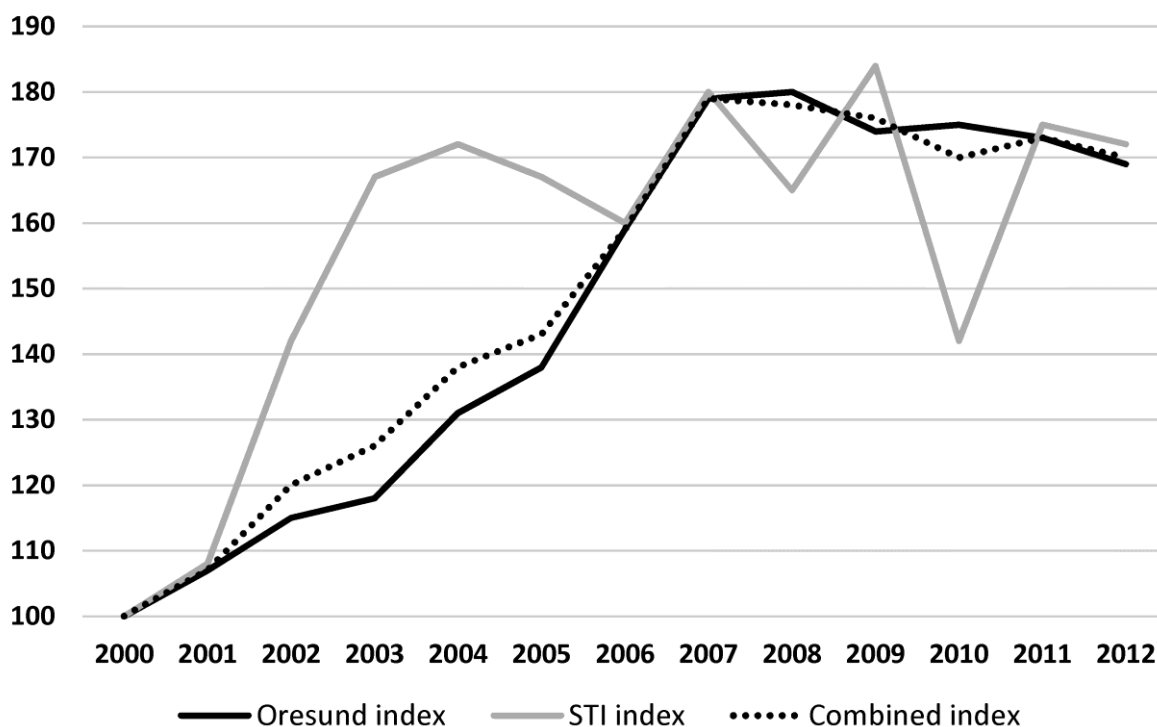
Figure 1 presents the individual basic, comparable and adjusted scores for co-patents, co-publications and collaborative R&D projects. This demonstrates the importance of adjusting the basic indices by taking into account comparable ones. Otherwise, there is the problem that general (national) trends could be confused with the actual effects of integration (Lundquist & Winther, 2006): for example, if there is a steady year-to-year increase in a comparable index, a corresponding increase in a basic index could mostly be explained by this general trend, hence the need to adjust it. As a general observation, there has been an almost steady increase in cross-border co-publications, while there has been more year-to-year variation in cross-border co-patents and collaborative R&D projects. This variation, which is in line with fluctuations in the sub-categories of the Index (Öresundskomiteen, 2015), will influence the composite STI index, as discussed below.

Figure 1: The basic, comparable and adjusted scores for cross-border A) co-patents, B) co-publications and C) collaborative R&D projects in the Oresund region (2000 = 100).



In general, the composite STI index follows the Index, but with some notable exceptions (Figure 2). It seems that the integration effects, which were (likely) associated with the opening of the Bridge, were more straightforwardly transferred to STI than to overall cross-border regional integration. According to Hansen (2013) this is partly due to the supporting policy measures which facilitated STI integration in the Oresund region. There was already a significant upsurge in the composite STI index in 2001–2003, while overall cross-border regional integration moved more steadily towards its peak in 2008 (index score = 180). The composite STI index reached its peak a year later in 2009 (index score = 184). The combined index peaks in 2007 with a value of 179. After 2008–2009, the Index and the composite STI index declined to 169 and 172 respectively in 2012, while the combined index decreased to 170 in the same period. Adding the composite STI index into the picture shows that when cross-border innovation collaboration is taken into account, the process of integration in the Oresund region seems to accelerate somewhat more quickly in the early years (2001–2003) of the observation period, but then slows down to follow the general pattern of the Index. This has raised concerns that the integration of the Oresund region could be in crisis, although the most recent data and forecasts provide cautious signs that integration is likely to start deepening in the near future after years of downturn and stagnation (Öresundskomiteen, 2015). However, as variations in the STI sub-category are naturally more erratic (due to the smaller number of variables) than the Index year-to-year comparisons of such a short time-period as the one covered by the Index might not do justice to actual integration processes within the CBR. Therefore, it might be more apt to state that the STI sub-category generally follows the developments of the Index.

Figure 1: The Oresund integration index (Öresundskomiteen, 2013) and the composite STI index compared (2000 = 100).



Additionally, there was a sudden year-long downward surge in the composite STI index in 2010. It would be tempting to consider that this indicates a definite sign of weaknesses in the

methodology inherent to composite indicators, or is the result of random variations in the dataset. However, the downward surge was mainly caused by the record-low number of new cross-border collaborative R&D projects at the same time as there was a relative slowdown in cross-border co-patents and a (gradual) recovery from a decrease in the number of cross-border co-publications a year earlier (Figure 2). Therefore, 2010 really was a period of exceptionally low cross-border STI collaboration rather than an anomaly resulting from methodological or data issues. Furthermore, the fact that the most recent decrease in the number of cross-border co-patents in 2012 is not obviously evident in the composite STI index supports the notion (Paas & Poltimäe, 2012) that utilising composite indicators is a way of overcoming some of the limitations of using single indicators. Firstly, this drop was also an EU-wide trend (Eurostat, 2015) and, thus, was already absorbed, to a certain degree, by adjusting the index. Secondly, since individual indicators can be prone to arbitrary year-to-year variations, the weighting scheme ensured that this abrupt downturn did not dramatically alter the entire STI index.

FUTURE RESEARCH DIRECTIONS AND CONCLUSION

Earlier studies on cross-border regional integration processes have largely neglected STI integration mainly due to not so much the lack of STI data per se, but the laborious nature of the data collection. Moreover, even the most notable example of a composite indicator approach to measuring cross-border regional integration, i.e. the Oresund integration index, lacks a sub-category for STI. This is regrettable since, by omitting innovation and knowledge flows, the Index fails to take into account one of the most important drivers of regional economic growth. Therefore, this paper has introduced a new composite indicator (sub-category) in order to measure STI integration in the case of the Oresund region; this was compiled from patent, publication and collaborative R&D project data.

The findings show that the revised composite indicator is relatively effective in depicting the process of integration *vis-à-vis* STI indicators, as well as when it is compared to the existing Index. Accordingly, while not entirely consistent with earlier industry-specific case studies (Moodysson & Jonsson, 2007; Hansen, 2013) showing only limited rates of cross-border innovation collaboration in the Oresund region, the findings of this paper clearly show that in terms of STI the Oresund region has become more integrated after the opening of the Bridge. However, it has to be kept in mind that, due to data availability issues, the suggested indicators (patents, publications and R&D projects) depict innovation in a rather narrow “STI” mode. A broader view, including also the “Doing, Using and Interacting” mode of innovation (Jensen, Johnson, Lorenz, & Lundvall, 2007), would require other indicators that are more challenging in comparable cross-border contexts.

In terms of the implications for other CBRs, earlier papers on the Oresund region suggest that policy should focus initially on the strongest industries which are present on both sides of the border and that there is a need to set up organizations to facilitate cross-border collaboration in a similar fashion to the Medicon Valley Alliance, an existing example of a successful platform for this kind of development (e.g. Hansen, 2013). The findings of this paper tentatively support that these policies have worked well, as shown by the intensified STI integration in the Oresund region. At a later stage, these industries could provide important role models for further integration in other sectors and segments of the economy (Lundquist & Trippel, 2009).

To conclude, while the data collection processes for the composite STI index developed in this paper are somewhat laborious, they are sufficiently straightforward to be updated relatively easily on a regular bases in the case of the Oresund region, and to be applied in other CBRs. The latter would facilitate conducting comparable cross-country analyses for the benefit of the empirical CBRIS literature.

REFERENCES

- BAK Basel Economics (2008). *Top Technology Region Benchmarking Report 2008*. Basel.
- Carvalho, N., Carvalho, L. & Nunes, S. (2015). A methodology to measure innovation in European Union. *International Journal of Innovation and Regional Development*, 6, 159-180.
- Coenen, L., Moodysson, J. & Asheim, B. (2004). Nodes, networks and proximities. *European Planning Studies*, 12, 1003-1018.
- Decoville, A., Durand, F., Sohn, C. & Walther, O. (2013). Comparing cross-border metropolitan integration in Europe. *Journal of Borderlands Studies*, 28, 221-237.
- Eurostat (2015). *Statistics explained: Patent statistics*. Retrieved November 25, 2015 from: http://ec.europa.eu/eurostat/statistics-explained/index.php/Patent_statistics
- Hagedoorn, J. & Cloudt, M. (2003). Measuring innovative performance. *Research Policy*, 32, 1365-1379.
- Hansen, T. (2013). Bridging regional innovation. *Geografisk Tidsskrift*, 113, 25-38.
- Hansen, T. & Hansen, R. (2006). Integration of the scientific community as exemplified by the biotech sector. *GeoJournal*, 67, 241-252.
- Jensen, M., Johnson, B., Lorenz, E. & Lundvall, B-Å. (2007). Forms of knowledge and modes of innovation. *Research Policy*, 36, 680-693.
- Lundquist, K-J. & Trippel, M. (2009). Towards cross-border innovation spaces. *RUW Discussion Papers*, 2009(5).
- Lundquist, K-J. & Winther, L. (2006). The interspace between Denmark and Sweden. *Geografisk Tidsskrift*, 106, 115-129.
- Makkonen, T. & van der Have, R.P. (2013). Benchmarking regional innovation performance. *Scientometrics*, 94, 247-262.
- Moodysson, J. & Jonsson, O. (2007). Knowledge collaboration and proximity. *European Urban and Regional Studies*, 14, 115-131.
- Moon, H-S. & Lee, J-D. (2005). A fuzzy set theory approach to national composite S&T indices. *Scientometrics*, 64, 67-83.

Nauwelaers, C., Maguire, K. & Ajmone Marsan, G. (2013). The case of Oresund. *OECD Regional Development Working Papers*, 2013(21).

OECD (2013). *Regions and Innovation: Collaborating across Borders*. Paris.

Örestat (2015). *Öresundsdatabasen*. Retrieved November 25, 2015, from:
<http://www.orestat.se/en/oresundsdatabasen-engelsk>

Öresundskomiteen (2013). *The Oresund Integration Index*. Copenhagen.

Öresundskomiteen (2015). *Öresund Integrationsindeks*. Copenhagen.

Paas, T. & Poltimäe, H. (2012). Consistency between innovation indicators and national innovation performance. *Eastern Journal of European Studies*, 3, 101-121.

Schmidt, T. D. (2005). Cross-border regional enlargement in Øresund. *GeoJournal*, 64, 249-258.



From emerging country to a leading role in the scientific and technological field? Analysis of the internationalization of Brazil¹

Cláudia Daniele de Souza^{*}, Daniela De Filippo^{**} and Elías Sanz-Casado^{***}

^{*}csouza@bib.uc3m.es; ^{**}dfilippo@bib.uc3m.es; ^{***}elias@bib.uc3m.es

Laboratory for Metric Information Studies (LEMI), Associated Unit of the Institute of Philosophy IFS-CSIC; Research Institute for Higher Education and Science (INAECU), University Carlos III of Madrid (UC3M), Calle Madrid, n. 126, Getafe-Madrid, 28038 (Spain)

ABSTRACT

Since 2002 the term BRIC was created to refer to a group of emerging countries with a growing economy, Brazil has been getting even more recognition in the international sphere. This date can be considered as a turning point for analyzing the process by which this country has passed (Bernal Meza, 2015). Overall it is possible to observe, first, an accelerated economic growth period followed very closely by a scientific and technological significant growth stage (Glanzel, Leta, Thijs, 2006). This second stage has been encouraged by numerous policies which main objectives were: promote the expansion of higher education, promote the internationalization of Research, Development and Innovation (R&D&I) and increase the quality and the visibility of Science and Technology (S&T).

The results of these policies can be traced by analyzing, for example, the evolution of the international scientific production. Thus, while in the 1980s Brazil was ranked 29th on the world by Web of Science (WoS) publications, was 24th in 1990, 17th in 2000 and 14th in 2005. Brazil has grown from 11 journals indexed in Wos (1997) to 125 (2014) and the number of public universities has also increased markedly, from 62 in 1980 to 103 in 2014.

There is no doubt that Brazil has already made the leap to the world stage. Their presence in different international alliances as the BRIC group (Brazil, Russia, India, China), the IBSA trilateral dialogue forum (India, Brazil, South Africa) and G20 that represents the interests of twenty developing countries at to the World Trade Organization (WTO) proves this. Furthermore, Brazil was one of the six main negotiators of the WTO Doha round, is involved in the dialogue between G-8 and aspires to obtain a permanent seat on the United Nations Security Council.

However, to what extent the country is integrated and recognized as a prestigious partner in the scientific landscape is still waiting to be explored.

Given the previous considerations, the aim of this work is to detect if, in recent years, there have only been a quantitative growth of scientific and technological activities, or if Brazil has also acquired international recognition and prestige from their scientific peers.

In this study we propose the analysis of the process by which Brazil has passed from being a "peripheral" country in science and technology (Kreimer, 2006; Beigel, 2013) to acquire a significant role in international S&T sphere. To study this internationalization process two stages are considered: the first one related to growth and the second the obtainment of recognition.

¹ This work was supported by Brazilian Federal Agency for the Support and Evaluation of Graduate Education (CAPES Brazil) with a full PhD scholarship, process. 0846-13-9.

INTRODUCTION

Since in 2002 the term BRICS² was created to refer to a group of emerging countries with a growing economy, Brazil has been getting even more recognition in the international sphere (GRATIUS, 2008). This date can be considered as a turning point for analyzing the process by which this country has passed (Bernal Meza, 2015). This significant increase is evident in the annual growth rates of Gross Domestic Product (GDP), which have been of 3.6% in 2005-2014, compared with 2.5% worldwide (UNCTAD Statistics 2015).

Certainly this growth is also often linked with Research and Development (R&D). So, generally it is possible to observe, first, an accelerated economic growth period followed very closely by a scientific and technological significant growth stage, once these investments lead to a growth in scientific outputs (Glanzel, Leta & Thijs, 2006). In the case of Brazil, this second stage has been encouraged by numerous policies which main objectives were: promote the expansion of higher education, the internationalization of R&D and increase the quality and the visibility of Science and Technology (S&T).

The results of these policies can be traced by analyzing the evolution of the international scientific production. Thus, while in the 1980s Brazil was ranked 29th on the world by Web of Science (WoS) publications, was 17th in 2000 and 14th in 2015. This ascension in the global scenario may result from the internationalization of national science and/or national journals. From 2005, brazilian journals have demonstrated progressive indexing in WoS (Testa, 2011), coming to 116 journals indexed in 2014 and promoting the increase of brazilian production in this database. Other indicators as the number of university (from 62 in 1980 to 103 in 2014) show the mentioned growth in academic sphere.

There is no doubt that Brazil had a lower level in the past, but in the recent 15 years, it achieved an amazing high speed of development especially in S&T, and already made the leap to the world stage (Myers, 2011). Their economic weights, as well as their scientific and technological production, are growing steadily.

Likewise the presence of Brazil in different international alliances as the BRICS countries, the IBSA trilateral dialogue forum (India, Brazil, South Africa) and G20 demonstrates the interest of the country to be an important role in the international arena.

Previous studies have analyzed the growing scientific activity in Brazil from the point of view bibliometric in the context of the BRICS countries (Kumar & Asheulova, 2011). Their rapid growth in scientific production and its impact has also been approached by authors like Mauleon & De Filippo (2013); Bornmann, Wagner & Leydesdorff (2015). Other study conducted in particularly scientific activity in Brazil in recent years investigates the relationship between type of university, numbers of degree program offered, faculty members, and papers (Velloso, Lannes & De Meis, 2004); and another analyzed the main instruments to become a global and regional power (Gratius, 2008).

² BRIC: Brazil, Russia, India and China. In 2010, South Africa became an official member of the group, hence BRICS.

However, to what extent Brazil is integrated and recognized as a prestigious partner in the scientific landscape is still waiting to be explored. In this line, given the previous considerations, the aim of this work is to detect if, in recent years, there have only been a quantitative growth of scientific and technological activities, or if Brazil has also acquired international recognition and prestige from their scientific peers.

SOURCES AND METHODOLOGY

In this study we propose the analysis of the process by which Brazil has passed from being a “peripheral” country in science and technology (Kreimer, 2006) to acquire a significant role in international S&T sphere. To study this internationalization process two stages are considered: the first one related to growth and the second the obtainment of recognition. The study of these stages has been operationalized through the dimensions and indicators presented in Table 1.

Table 1. Dimensions, indicators and sources

Dimensions	Indicators	Sources
GROWTH	Number of researchers; researchers per 1000 inhabitants; R&D spending per researcher; S&T spending in relation to GDP	RICYT
	Number of doctoral theses defended	Capes
	Number of scholarship for international mobility	
	Number of public Universities	Ministry of Education (Brazil)
	Number of WoS journals	Journal Citation Report (JCR)
	Number of WoS publications	Thomson Reuters
RECOGNITION	% of documents in international collaboration	Web of Science
	% of documents in the first quartile (Q1)	
	Publications impact and international excellence (Highly Cited Papers) Number of Citations, Citations per paper; % Papers Cited Citations by subject area	Essential Science Indicators
	Presence in international rankings of universities	ARWU, THE, QS
	Participation in European projects of Framework Programme	CORDIS

This study has been conducted following the next steps:

- Obtaining of S&T activity indicators for Brazil from different official sources;
- Recovery and treatment of brazilian scientific publications indexed in Web of Science (WoS) through bibliometric indicators of: activity, collaboration, impact and visibility;
- Checking of the *Essential Science Indicators* to detect the Brazil position in the world and the most cited Brazilian institutions;
- Selection of the *Highly Cited Paper* (most cited publications in 10 years);
- Analysis of international rankings of universities: Academic Ranking of World Universities (ARWU), QS World University Rankings (QS) and Times Higher Education (THE) to obtain positioning indicators of Brazilian public universities;
- Selection and download of data on Brazil's participation in European Framework Programme Projects

RESULTS

Indicators that show different aspects of the evolution of S&T of Brazil are presented below.

GROWTH INDICATORS

Table 2 shows the growth of scientific activity in Brazil, considering specific years of the last decades. The data show that the number of researchers has doubled in 15 years, a situation that has not been observed in any other Latin American country. Compared to all the countries of Latin America, Brazil ranks third in researchers per thousand inhabitants as in R&D spending per researcher in thousands of dollars, being behind only Argentina and Costa Rica throughout the period (RICYT, 2016)

It is possible to also be observed that the articles published in WoS increased from 0.30% of the total database in the early 1980s reaching 2.35% of world production in 2014.

Table 2. Growth indicators considering specific years in recent decades

Indicators	1980	1990	2000	2014
Number of researchers	--	--	96.916	170.209
Researchers per thousand inhabitants	--	--	1.54 ³	2.51 ⁴
R&D spending per researcher in thousands of dollars	--	--	52.08	109.12
Number of theses defended per year	1.005	1.410	5.318	15.287
Number of public universities	62	79	97	103
Number of scholarship for international mobility	--	1.867	2.492	26.218
Number of journals in JCR (WoS)	--	11 ⁵	19	116
Number of WoS publications	2.253	4.064	13.558	47.928
% docs Brasil en WoS	0.30%	0.46%	1.12%	2.35%

RECOGNITION INDICATORS

Among the indicators related to scientific production is observed that the percentage of documents in international collaboration has increased from 14% in 1980 to 32% in 2014. This evidence that Brazil has become an important partner for a growing number of countries. United States, United Kingdom, Spain, France and Germany are considered amongst its main partners.

Regarding the documents in the first quartile (Q1), in recent years, although the absolute values have increased, the percentage has declined (Table 3).

Table 3. Recognition indicators of scientific production considering specific years in last decades

Indicators	1980	1990	2000	2014
% WoS documents in international collaboration	14.34%	25.91%	31.34%	32.28%
Number of partners in collaboration	48	57	110	173
% of documents in Q1	--	--	50.00%	41.24%

³ RICYT, 2001

⁴ RICYT, 2010

⁵ JCR 1997

Impact indicators measured by citations received show that Brazil is currently ranked 17th in the world (Thomson Reuters, 2016). However, there are areas in which improvement this position such as *Agricultural Science* (8th position); *Plant & Animal Science* (12th); *Environmental/Ecology* (15th), *Immunology* (15th). Moreover, in the last decade the number of citations per document has increased in all areas. Table 4 shows the distribution of received citations by subject area. It can be observed that while Brazil's contribution to total world production is 2.35%, there are areas such as Plant & Animal Science or Agricultural Sciences in which these percentages far outweigh the average global impact.

Regarding the percentage of cited documents, Brazil has similar values to the world average, although highlighted in *Chemistry* and *Multidisciplinary Science*. Concerning the number of citations per document the most important area is *Molecular Biology* (Table 4).

Table 4. Distribution of citations by subject area in Brazil and comparison to the world average

Áreas <i>Essential Science Indicators</i>	N.Papers WoS Brazil	% Docs Brazil / World	Citations Brazil	% Papers Cited Brazil	% Papers Cited World	Citas/paper Brazil	Citas/paper en el World
Clinical Medicine	95.826	2.04%	655.937	58.97%	52.13%	6.85	7.81
Chemistry	31.147	1.69%	297.331	81.11%	73.56%	9.55	12.52
Physics	25.764	2.17%	283.728	82.75%	80.92%	11.01	11.94
Plant & Animal Science	45.891	5.50%	223.895	66.72%	68.70%	4.88	13.28
Biology & Biochemistry	22.526	2.30%	175.306	65.21%	66.14%	7.78	21.82
Agricultural Sciences	37.967	8.97%	170.893	65.46%	69.97%	4.5	12.37
Neuroscience & Behavior	15.905	2.01%	141.194	66.60%	59.79%	8.88	7.48
Environment/Ecology	12.319	2.92%	130.495	76.14%	80.77%	10.59	10.45
Engineering	16.845	1.39%	115.894	70.92%	67.96%	6.88	8.28
Molecular Biology & Genetics	10.200	1.89%	110.623	72.59%	74.09%	10.85	4.71
Pharmacology & Toxicology	13.497	2.55%	103.476	67.96%	62.30%	7.67	13.21
Immunology	10.184	2.64%	97.361	66.03%	63.08%	9.56	11.35
Materials Science	12.264	1.59%	85.511	71.14%	73.66%	6.97	13.56
Social Sciences, general	19.655	1.51%	73.792	51.58%	47.12%	3.75	9.57
Microbiology	7.432	3.55%	72.767	82.66%	82.15%	9.79	9.02
Geosciences	7.336	1.59%	71.946	76.72%	75.51%	9.81	8.24
Space Science	3.569	2.39%	54.283	85.79%	86.33%	15.21	15.41
Psychiatry/Psychology	7.141	1.28%	44.423	52.96%	57.66%	6.22	6.34
Mathematics	7.972	1.96%	31.310	65.38%	65.29%	3.93	18.73
Computer Science	7.138	1.41%	30.799	60.02%	63.65%	4.31	7.57
Economics & Business	2.469	0.83%	9.998	50.99%	61.40%	4.05	4.56
Multidisciplinary	518	0.66%	4.117	59.46%	30.93%	7.95	4.75

Information about '*Highly cited papers*' in WoS over the past 10 years shows, among the first 500 items in the world, by received citations, the best position of Brazil is a document published by the Clinical Hospital of the University of São Paulo, with 3.297 citations

received (ranked 17th in the world) in *Clinical Medicine*. This area, with *Physics* and *Environmental/Ecology* shows the greatest number of ‘*Highly cited papers*’ with the best positions between 15th and 19th position of the world.

Considering visibility indicators in international rankings in ARWU Brazil has evolved from 4 universities in 2004 to 6 in 2015 in the top 500 in the world, while other Latin American countries does not exceed by 3 universities in this ranking. In QS ranking until 2010, there are not none university among the top 500, while in 2015 there are 7. Currently in THE there are 2 universities (Table 5).

Table 5. Presence of Latin American universities in international universities rankings

Latin American universities in the rankings	ARWU			QS			THE ⁶	
	2004	2010	2015	2004	2010	2015	2010	2015
Brasil	4	6	6	0	3	7	0	2
México	3	2	2	1	3	3	0	2
Argentina	1	1	1	0	2	6	0	1
Chile	1	2	2	0	2	3	0	2
Colombia	0	0	0	0	0	4	0	0

Table 6 shows Brazilian universities present in the three main international rankings: ARWU, QS and THE.

Table 6. Presence and position of Brazilian universities in international rankings

Position in the Ranking ARWU 2015	Brazilian universities	Position in the Ranking QS 2015	Brazilian universities	Position in the Ranking THE 2015	Brazilian universities
101-150	USP	143	USP	201-250	USP
301-400	UNICAMP	195	UNICAMP	351-400	UNICAMP
301-400	UNESP	323	UFRJ		
301-400	UFRJ	451-460	UFRGS		
401-500	UFMG	481-490	UNESP		
401-500	UFRGS	491-500	UnB		
		491-500	UNIFESP		

Concerning the presence as a partner in European projects (Table 7), the participation of Brazil has grown to get to 17 projects in the Seventh Framework Programme, which shows the interest and ability to be part of networks of high international quality.

⁶ 'THE' ranking only have data available in the period 2011-2016.

Table 7. Participation of Brazil in European Framework Programme Projects

Period	Programme	Number of projects with Brazilian participation	Number of projects in the world
1998-2002	FP5	51	17.202
2002-2006	FP6	99	10.103
2007-2013	FP7	170	25.594
2014-2020	H2020	17	6.435

CONCLUSIONS

In recent years the growth of emerging countries such as the BRICS has been remarkable in various fields and their contribution to economic activities is increasing. In fact in 2014, BRICS economies generated more than 20% of the world's GDP (UNCTAD Statistics 2015). Considering scientific publication, from 1990 to the present time a remarkable growth is observed. Its scientific production in WoS increased from 7% of the world up to 17% in 2010. Of the four, China and Brazil present the greater growth in the period (Mauleon & De Filippo, 2013).

In the case of Brazil, their rapid development has been remarkable in comparison with its immediate environment, the Latin American region. Some indicators that prove it are investing in R&D that has gone from 1,01% of GDP in 1990 to 1.24% in 2013, while the entire region did not exceed 0,8% in those years. Growth in the number of researchers is also notable going from 1.54% of the economically active population in 2001 to 2.51% in 2010, while the average in the region did not exceed 1.5%. If one considers the number of doctors in 1990 Brazil contributed 59% of the region, but two decades later, the percentage rose to 74%. In terms scientific production, Brazil's growth has been much higher than in Latin America, contributing 35% in 1990 compared to 55% in 2013 (RICYT, 2016).

Such as mentioned by some authors, the growing of scientific output is to some extent driven by human resources mobility and international collaboration. According to Abramo et al (2011) this phenomenon can be linked to a number of factors, including the implementation of specific policies favoring research collaboration, at various levels. In the case of Brazil different programs implemented in the past decade undoubtedly have contributed in this line. Among them the Brazilian Program "Science Without Borders" that it is a large scale scholarship program promoted by the Ministry of Education and the Ministry of Science and Technology through their respective funding agencies CAPES and CNPq. This program has been possible strengthen and expand the initiatives of science and technology, innovation and competitiveness through international mobility of undergraduate and graduate students and researchers. Then, the return of these researchers to their homelands constitutes a strong transfer of science and technology, in addition to the fact that they typically maintain collaborative ties with their host institutions (Science without Borders Program, 2016).

The increase in international scientific production in Brazil can also be attributed to the efforts that the country has been undertaking at the national level. Investments in formation and development of researchers have stimulated scientific activity and also activities related to the publishing of national journals. Production has been advancing increasingly to foreign journals. In this line, the National Plan for Postgraduate studies was designed as a route for

accelerating the training of human resources suitable to supply the urgent need for qualified personnel capable of improving the quality of teaching and strengthening science and technology activities in Brazil, contributed a lot to the quality of its scientific production (Guimarães & Human, 1995).

But not only has increased production, also grown the presence of Brazilian universities in international rankings, the notable increase in the number of journals indexed in international databases, the percentages of international collaboration, the participation in European projects and the increasing number of citations, show clearly an increasing inclusion in the international scientific community.

All these indications suggest a witnessing the transition from a developing economy to a “knowledge economies”. As it mentioned in a World Bank study (Dahlman & Aubert, 2001), in the contemporary world, rapidly developing economies tend to be those in which economic growth depends increasingly on the creation, acquisition, distribution and use of knowledge. A fundamental prerequisite for this type of economy is human capital consolidated, particularly, a population with a fairly high general level of education and a significant proportion of people with a higher education in science and technology.

Despite this evident increase, there are still aspects to improve like the quality of production (measured by the percentage of publications in Q1) and the impact received by citations. In this line some remarkable scientific areas have been identified but the overall average production has yet to make the leap to a higher quality. Undoubtedly an aspect that can be achieved if investment in R&D remains a priority for the country.

REFERENCES

- Abramo, G., D’Angelo, C. A., Solazzi, M. (2010). The relationship between scientists’ research performance and the degree of internationalization of their research. *Scientometrics*, 86(3), 629-643.
- Bernal Meza, R. (2015) La inserción internacional de Brasil: el papel de BRICS y de la región. *Universum*, 30(2), 17-35.
- Bornmann, L.; Wagner, C.; Leydesdorff, L. (2015) BRICS countries and scientific excellence: A bibliometric analysis of most frequently cited papers. *Journal of the Association for Information Science and Technology*, 66(7), p.1507–1513.
- Dahlman, C.; Aubert, J. E. *China and the Knowledge Economy. Seizing the 21st Century*. Washington, DC: World Bank, WDI Development Studies, 2001.
- Gratius, S. (2008) Brazil emerges as a regional and global power. *Quórum: revista de pensamiento iberoamericano*, 22, p.135-146.
- Glanzel, W., Leta, J., & Thijs, B. (2006). Science in Brazil. Part 1: A macro-level comparative study. *Scientometrics*, 67(1), 67-86.

Guimarães, J., Humann, M. (1995). Training of Human-Resources in Science and Technology in Brazil - the Importance of a Vigorous Postgraduate Program and Its Impact on the Development of the Country. *Scientometrics*, 34(1), 101-119.

Kreimer, P. (2006) ¿Dependientes o integrados?: La ciencia latinoamericana y la nueva división internacional del trabajo. *Nómadas*, 24, p. 199–212.

Kumar, N; Asheulova, N. (2011) Comparative analysis of scientific output of BRIC countries. *Annals of Library and Information Studies*, 58, p. 228-236.

Myers D. (2011) Selling to the BRIC: Brazil - an awakening tiger. *Learn Publ.*, 24(4):261–5.

Mauleon E, De Filippo D. (2013) Are the BRIC and MITS countries improving their presence in the International Science? In: Gorraiz J, Schiebel E, Gumpenberger C, Horlesberger M, Moed H. *14th International Society of Scientometrics and Informetrics Conference (ISSI)*. Leuven: Int Soc Scientometrics & Informetrics-Issi. p. 1868–1870.

Red de Indicadores de Ciencia y Tecnología Iberoamericana e Interamericana - RICYT (2016). *Indicadores comparativos*. Retrieved July 14, 2016 from: <http://www.ricyt.org/indicadores>

Science without Borders Program (2016). Retrieved July 13, 2016 from: <http://www.cienciasemfronteiras.gov.br/web/csf-eng/>.

Testa, J. *The globalization of Web of Science*. Thomson Reuters Expert Essays. 2011. Retrieved July 11, 2016 from: http://wokinfo.com/products_tools/multidisciplinary/webofscience/contentexp/expansionessay/

UNCTAD Statistics. (2015). *Real GDP growth rates, total and per capita, annual, 1970–2014*. Retrieved July 13, 2016 from: <http://unctadstat.unctad.org/wds/TableViewer/tableView.aspx?ReportId=109>

Velloso A, Lannes D, De Meis L. Concentration of science in Brazilian governmental universities. *Scientometrics*, 61(2):207–20.



Measuring internationality without bias against the periphery¹

Valeria Aman*

*aman@dzhw.eu

DZHW– German Centre for Higher Education Research and Science Studies
Schützenstraße 6a. 10117 Berlin (Germany)

ABSTRACT

This paper explores the degree of internationality in the context of academic publishing. Therefore, an Internationality indicator based on the Gini-Simpson-Index was created. It captures not only the composition of countries authoring but also the countries citing a journal. The indicator is applied to study the internationality score of center and peripheral countries among six research fields. Unlike previous indicators measuring internationality, the indicator proposed in this paper is not biased towards peripheral countries. Our results show that internationality is not uniform and some areas like natural sciences are more international than other areas. This contribution demonstrates that the Internationality indicator proposed is not biased against peripheral countries, so that the true internationality of publications from these countries is captured.

INTRODUCTION

The purpose of the study is to explore internationality in the context of academic publishing and to propose a new indicator. Literally, “international” means between countries. With regard to authorship a journal publishing authors from two neighbouring countries would be classified international as well as a journal publishing authors from all over the world, i.e., multinational. Buela-Casal et al. (2005) argue that scientometrists use the terms “international” and “multinational” interchangeably, whereby an “international journal” is expected to reflect a global perspective rather than fulfilling simply the literal definition. Accordingly, a quantitative index to express the degree of internationality is needed.

Different criteria and approaches have been proposed to assess journal internationality. Zitt and Bassecouard (1998) studied the internationality of journals in Earth & Space and Applied Biology by comparing the distribution of authoring and citing countries with average profiles of a discipline. Journals are considered as international if their spectrum is close to the average country distribution in their discipline. Zitt and Bassecouard mentioned the bias of their indicator “to produce high rates of internationalization for US journals” (1998, p. 267), because the world reference is shaped by these journals.

In another study, Zitt and Bassecouard (1999) proposed an internationalization measure at the country-level that considers the relation of national-oriented and international-oriented journal publications of a country.

In addition to publication distributions, Christensen and Ingwersen (1996) proposed citation distributions to measure the internationality of main journals in Library and Information

¹ The author gratefully acknowledges use of the database and services of the *Competence Centre for Bibliometrics*, funded by the BMBF (01PQ13001).

Science. Several studies on internationality from a micro-perspective (journal samples) followed (see Wormell, 1998; Uzun, 2004; Sin, 2011). Perakakis et al. (2005) evaluated the pattern of multinational distribution of authors by calculating the Gini-Coefficient of four psychology journals. Due to the ambiguity of their results, they suggested a new journal internationality index to be constructed. Studying the internationality of four psychology journals and identifying criteria amenable to quantitative analysis, Buena-Casal et al. (2005) noted that “a single quantitative measure of internationality [...] would be an indispensable tool in the hands of authors, readers, editors, publishers and generally anyone interested in the evaluation of journals” (p. 45).

In response to previous work, the approach presented in this paper not only considers the multinational distribution of authors, but also of users. Since bibliometrics can only capture the geographical distribution of citers (and not readers), it is assumed that the geographical distribution of citers reflects the multinational distribution of users.

The internationality indicator proposed in the following is applied to study the internationality of center and peripheral countries. According to Englander and Smith (2013) “peripheral” is a contested but helpful term to refer to nations that are economically disadvantaged in terms of investments in research and development relative to the “center” regions of the world and where English is not the dominant language (p. 232). Arunachalam (1992) reported that science on the periphery is characterised by insularity, i.e., a lack of contact with international science and where peripheral scientists seldom have the chance to work in newly emerging research fields. Journals from peripheral countries are often of poor quality and the elite of local scientists prefer to publish in international journals (ibid.).

Since journals with few contributions from center countries are penalized in their degree of internationality in previous approaches (Zitt and Bassecouard, 1998; Zitt and Bassecouard, 1999; Buena-Casal et al. 2005), the indicator proposed rates any geographically heterogeneous pattern of authors and users as international, regardless of the authoring or citing countries.

INTERNATIONALITY INDICATOR

In the following, the internationality indicator is introduced. First of all, the total publication counts have to be calculated. To reflect the diversity of readership, citations are included. A set of publications from a single country is considered more international if it reaches a foreign readership. Thus, P represents the sum of publication counts over all countries publishing in a publication set, whereas C is the sum of the citations of each country publishing in the publication set.

$$P = \sum_{Country} P_{country}$$

$$C = \sum_{Country} C_{country}$$

The indicator is inspired by the Gini-Simpson-Index (Simpson, 1949) that found its way into bibliometrics. Wang et al. (2015) used the Gini-Simpson-Index to study the association between the degree of interdisciplinarity and scientific impact. In our case, the Gini-Simpson-Index I_o can be interpreted as the probability of obtaining different countries when randomly choosing (with backplacing) two publications from the set of publications. The Gini-Simpson-Index I_o is calculated as one minus the sum of the squared publication shares of the countries:

$$I_o = 1 - \sum_{country} \frac{P_{country}^2}{P^2}.$$

The following statements hold:

$I_o = 1$ if and only if $P = 0$, and

$I_o = 0$ if $P_{country} = 0$ for all but one country.

The internationality indicator I can be interpreted as the probability that a randomly chosen publication and another randomly chosen publication or citation are from different countries. It is calculated as one minus the sum of the products of the publication shares and the *publication-citation shares* of each country, where the *publication-citation share* of a country is defined as the proportion of the countries' combined publications and citations in all publications and citations:

$$I = 1 - \sum_{country} \frac{P_{country}}{P} \cdot \frac{C_{country} + P_{country}}{C + P}.$$

The following statements hold:

$I = 1$ if and only if $P = 0$,

$I = I_o$ if $C = 0$, and

$I = 0$ if $P_{country} = 0$ and $C_{country} = 0$ for all but one country.

The indicator can take values between 0 and 1. Values close to 1 reflect internationality, whereas values close to 0 express nationality. The indicator does not differentiate between publications from neighbouring countries and from countries located in different regions of the world. Being cited is not a necessary condition for a high internationality score. The idea here is that citations are measured only if they are included in a citation database.

DATA AND METHOD

The data originate from an in-house database of *Thomson Reuters' Web of Science* (WoS), provided as tagged data in terms of a consortia license. The analyses of publications and citations are restricted to the *Science Citation Index-Expanded* (SCI-E), the *Social Sciences Citation Index* (SSCI) and the *Arts & Humanities Citation Index* (A&HCI). All journal publications of the document type *article*, *letter*, and *review* have been considered as source items. Citations received by these papers were determined for the publication period 2008-2011. A moving 4-year citation window has been applied, so that citations from the actual year of publication and the three following years are considered. This 4-year-citation window allows calculating citation rates up to 2014. The OECD Category scheme² was used that consists of two levels: six major fields and 42 minor disciplines. All 252 WoS *subject categories* are represented in the mapping. According to *Thomson Reuters'* all six OECD fields are represented in the final mapping, but not all forty-two OECD disciplines.³

Fractional counting was applied on the country level. Thus, a collaborative publication of four different countries attributes each of the countries a fourth (0.25) of the publication. Likewise, the citation counts are fractionally attributed, so that each country receives one fourth of a citation (0.25). Since a journal may belong to several *subject categories*, and thus OECD classes, these have been also fractionally assigned.

²<http://ipscience-help.thomsonreuters.com/incitesLive/globalComparisonsGroup/globalComparisons/subjAreaSchemesGroup/oeed.html>

³ The following 3 codes have been deliberately omitted by Thomson Reuters: 4.04 Agricultural Biotechnology 3.04 Health Biotechnology 3.04 Other Medical Sciences

RESULTS AND DISCUSSION

For reasons of comparison I calculated the internationality indicator proposed by Zitt and Bassecouard (1998) on the basis of the data as described and concluded that their indicator prefers center countries over peripheral countries. Their internationality score turns out to be low even if several countries contributed to a journal (e.g. Japan, Greece) in comparison to a journal that has only publications from the US. This is due to the fact that the coverage of journals in WoS affects the optimal distribution of a highly international journal (ibid). This does not reflect a fair indicator towards peripheral countries since WoS is biased towards US-journals and all the more back in 1998, when Zitt and Bassecouard published their results. The comparison of their internationality score and the one proposed here is illustrated with an example: According to their indicator, journals classified in WoS' subject categories "Literature, American" or "Law" have high internationality scores, whereas the indicator presented here reaches low internationality scores. To my understanding this is reasonable, since most of the authors contributing to a journal in *American Literature* are from the US. On the other hand, the majority of journals covered in the subject category *Law* are of US-origin. Thus, authors contributing to these journals are mostly from the US and moreover, *Law* is a field of national interest. In contrast to Zitt and Bassecouard (1998), the internationality score proposed in this paper leads to rather low scores, because the authorship and readership in journals assigned to "Literature, American" or "Law" do not reflect heterogeneous distributions.

A high internationality score can be reached through a high share of international co-authored papers or through contributions from various countries. A low internationality score can have its roots in journals where foreign researchers rarely publish or in the underlying discipline. Table 1 provides an overview of the 39 OECD disciplines ordered according to their degree of internationality.

Table 1. Overview of OECD disciplines and their Internationality score (I-Score) for journals publishing in 2008-2011 and a citation window of 4 years.

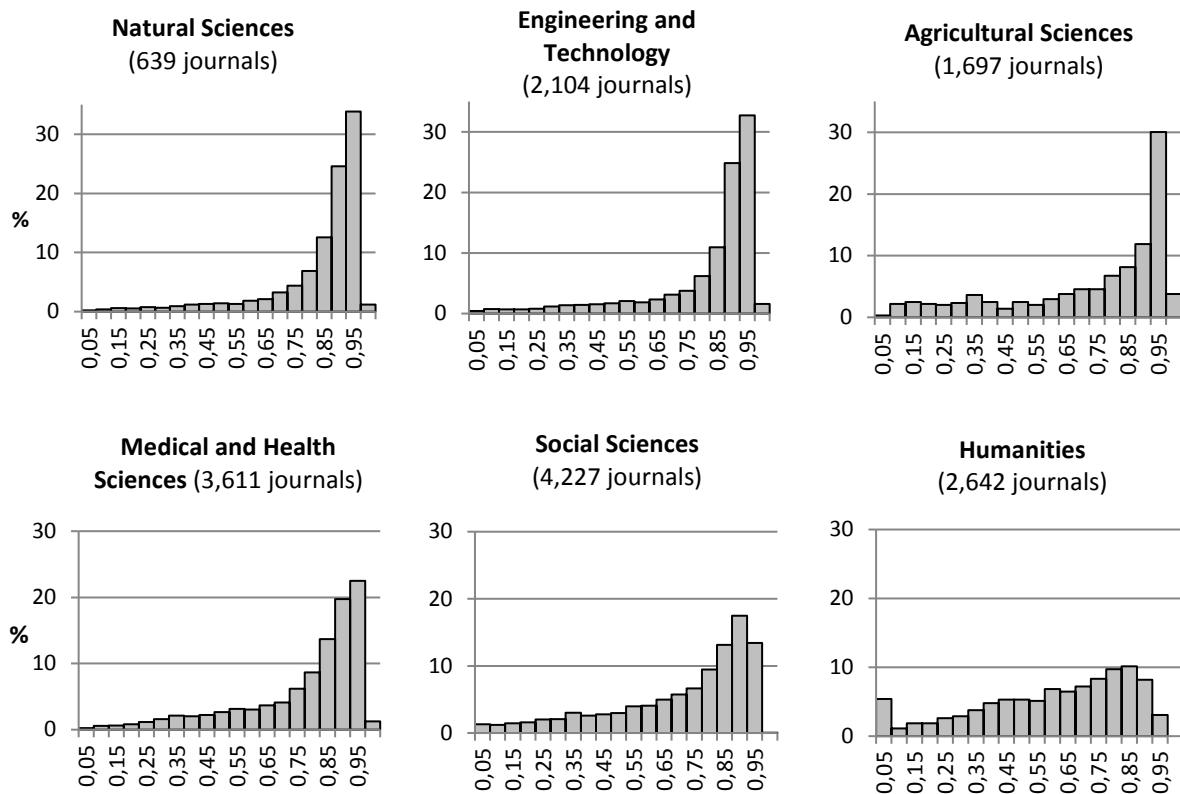
OECD Discipline	I- Score	OECD Discipline	I- Score
Industrial Biotechnology	0.882	Economics and business	0.795
Nano-technology	0.868	Clinical medicine	0.789
Environmental biotechnology	0.855	Other social sciences	0.768
Computer and information sciences	0.851	Agriculture, forestry, and fisheries	0.741
Mathematics	0.849	Health sciences	0.739
Physical sciences	0.847	Other natural sciences	0.739
Mechanical engineering	0.846	Psychology	0.728
Environmental engineering	0.844	Animal and dairy science	0.713
Civil engineering	0.841	Veterinary science	0.706
Biological sciences	0.840	Media and communications	0.691
Chemical engineering	0.838	Political Science	0.652
Chemical sciences	0.838	Sociology	0.651
Medical engineering	0.837	Philosophy, ethics and religion	0.647
Electrical engineering, electronical engineering, information engineering	0.836	Educational sciences	0.629
Other engineering and technologies	0.831	History and archaeology	0.606
Other agricultural sciences	0.829	Art (arts, history of arts, performing arts, music)	0.589
Earth and related environmental sciences	0.829	Languages and literature	0.572
Basic medicine	0.822	Other humanities	0.560
Materials engineering	0.802	Law	0.449
Social and economic geography	0.797		

The most international discipline is *Industrial Biotechnology*, followed by *Nano-technology*. National disciplines are *Law* and the humanities in general. Medical researchers tend to publish their different experiences in national journals, so that the score is rather low.

One could argue that journals included in WoS are *per se* international, since it is one of the inclusion criteria (see Testa, 2016). Nevertheless, the degree of internationality differs among disciplines. Figure 1 provides an overview of the skewness of the internationality score of journals among the OECD fields. The journal internationality score was calculated for the publication year 2011 and a citation window of four years.

The highest percentage of highly international journals is found in *Natural Sciences* and *Engineering and Technology*, both fields showing a similar distribution. In *Agricultural Sciences* the journal internationality scores are close to 1 (mostly journals on *Food Science*), next to rather national journals. *Social Sciences* and the *Humanities* reveal a high share of journals that are rather national than international. Note that the internationality indicator depends on the number of articles published in a year, which differs in *Natural Sciences* and the *Humanities*.

Figure 1. Distribution of journals according to their internationality score among the six OECD research fields. The number of journals covered is provided in parentheses.



Based on these six OECD fields the internationality indicator was processed for the 20 countries with the highest publication output in the 4-year period 2008 to 2011. The distribution of scientific output as reflected in WoS is skewed, so that the 20 countries listed in Table 2 account for 85% of the published journal literature in 2008-2011. The OECD fields are arranged in

accordance to their average internationality score. The countries are sorted in descending order, according to the average internationality score among all OECD fields.

Table 2. Overview of the average journal internationality score of the top-20 publishing countries in 2008-2011 according to OECD field.

Country	Natural Sciences	Engineering and Technology	Medical and Health Sciences	Agricultural Sciences	Social Sciences	Humanities
Netherlands	0.876	0.871	0.851	0.854	0.796	0.725
Taiwan	0.864	0.855	0.825	0.874	0.852	0.668
Italy	0.880	0.887	0.844	0.840	0.835	0.603
United Kingdom	0.869	0.874	0.838	0.843	0.776	0.658
Iran	0.842	0.849	0.745	0.790	0.852	0.719
Australia	0.859	0.873	0.790	0.806	0.731	0.663
Switzerland	0.867	0.870	0.824	0.697	0.779	0.679
Canada	0.855	0.870	0.799	0.793	0.740	0.596
Japan	0.825	0.771	0.822	0.675	0.779	0.741
Spain	0.884	0.872	0.728	0.867	0.665	0.580
France	0.872	0.870	0.738	0.804	0.746	0.564
China	0.735	0.756	0.832	0.854	0.795	0.589
South Korea	0.811	0.798	0.779	0.638	0.773	0.670
Germany	0.870	0.805	0.734	0.707	0.736	0.598
India	0.797	0.845	0.799	0.530	0.752	0.701
United States	0.836	0.835	0.753	0.748	0.615	0.520
Turkey	0.830	0.834	0.735	0.713	0.610	0.408
Poland	0.807	0.638	0.677	0.527	0.669	0.491
Brazil	0.762	0.815	0.667	0.311	0.605	0.466
Russia	0.564	0.526	0.409	0.485	0.390	0.282

In Table 2 we find high internationality scores among center countries and among small countries where no national scientific publishers exist (e.g. Taiwan). The Netherlands have very high internationality scores. Zitt and Bassecouard (1999) also noted that the Netherlands are outstanding “with the well-known home-published internationalized journals” (p. 681). It is striking that Iran as a rather peripheral country shows a highly international publication profile. This might result from the fact that influential researchers prefer to publish their results in English-language journals of center countries. On the opposite, the peripheral countries Turkey, Poland, Brazil and Russia show a rather nationally oriented publication profile.

Brazil’s low internationality score in *Agricultural Sciences* may be due to “disciplinary insularity”, a term used by Arunachalam (1992, p. 73). Thus, the majority of citations to publications from Brazil just as to scientists from Brazil publishing in international journals, are from the same national scientific community. The disciplinary insularity is also evident in Japanese *Engineering and Technology* publications. The relatively low score suggests that there is a strong national community publishing in this field.

Russia has a remarkably low internationality score in all of the OECD fields. This indicates that scientists from Russia depend on national journals, addressing a national community and covering national interests. To reach a wider international readership some peripheral journals from Russia are supported by the international publishing house Springer (Roitman, 2004).

CONCLUSION

Buela-Casal et al. (2005) noted that internationality remains a “mathematically fuzzy entity” and an “Internationality Index” has to be constructed from the combination of weighted criteria (p. 45). In this paper, an internationality indicator was proposed that builds upon the distribution of authoring and citing countries. The indicator is less biased towards peripheral countries as compared to previous approaches. One limitation of the indicator is the dependence on the journal coverage in the database under study. If for example Thomson Reuters adds or removes national-oriented journals, this affects the internationality score of countries in which the added or removed journals represent an important fraction of their output. WoS is arguably biased in favor of scientifically center countries and as long as it is used to assess the scientific performance of peripheral countries, it grasps only the contributions of these countries to international journals, which is often only a small share of their total scientific output. Nevertheless, scientists from peripheral countries publish their research in national journals. These journals serve to establish ties between research of interest to scientists in center and peripheral countries, whereas international journals increase the diversity of approaches and viewpoints on research. In conclusion, Buela-Casal et al. (2005) argue in their study that no single criterion alone is sufficient to capture the concept of internationality. Whereas the internationality indicator proposed in this paper focused on authoring and citation distributions, future research might consider the language of publications, editorial boards, and other potential criteria to study the internationality of publication sets.

REFERENCES

- Arunachalam, S. (1992). Peripherality in science : what should be done to help peripheral science get assimilated into mainstream science. In Rigas, A.; Gaillard, J. (ed.). *Science indicators for developing countries*. Paris : ORSTOM, 1992, 67-76.
- Buela-Casal, G., Perakakis, P., Taylor, M., & Checa, P. (2006). Measuring internationality: Reflections and perspectives on academic journals. *Scientometrics*, 67(1), 45–65. doi:10.1556/Scient.67.2006.1.4
- Christensen, F. H., & Ingwersen, P. (1996). Online citation analysis. *Scientometrics*, 37(1), 39–62. doi:10.1007/BF02093484
- Englander, K., & Smith, S. U. (2013). The role of policy in constructing the peripheral scientist in the era of globalization. *Language Policy*, 12(3), 231–250. doi:10.1007/s10993-012-9268-1
- Perakakis, P., Taylor, M., & Buela-Casal, G., Checa, P. (2005). A neuro-fuzzy system to calculate a journal internationality index. *Proc. CEDI (2005), Granada, Spain., 1*, 157–163. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.170.782&rep=rep1&type=pdf>
- Roitman, E.V. (2004). Development of the biomedical press in modern Russia. *The Write Stuff* 13,2: 46-47.

- Sin, S.-C. J. (2011). Longitudinal Trends in Internationalisation, Collaboration Types, and Citation Impact. *Journal of Library and Information Studies*, 9(1), 27–49.
- Testa, J. (2016). The Thomson Reuters Journal Selection Process. Retrieved from: <http://wokinfo.com/essays/journal-selection-process/>
- Uzun, A. (2004). Assessing internationality of scholarly journals through foreign authorship patterns: The case of major journals in information science, and scientometrics. *Scientometrics*, 61(3), 457–465. doi:10.1023/B:SCIE.0000045121.26810.35
- Wang, J., Thijs, B., & Glänzel, W. (2015). Interdisciplinarity and Impact: Distinct Effects of Variety, Balance, and Disparity. *PLOS ONE*, 10(5), e0127298. doi:10.1371/journal.pone.0127298
- Wormell, I. (1998). Informetric analysis of the international impact of scientific journals: How “international” are the international journals? *Journal of Documentation*, 54(5), 584–605. doi:10.1108/EUM0000000007182
- Zitt, M., & Bassecouard, E. (1998). Internationalization of scientific journals: A measurement based on publication and citation scope. *Scientometrics*, 41(1), 255–271.
- Zitt, M., & Bassecouard, E. (1999). Internationalization of communication a view on the evolution of scientific journals. *Scientometrics*, 46(3), 669–685. doi:10.1007/BF02459619



Indicators on measuring technology convergence worldwide¹

Chunjuan Luan*

*julielecj@163.com

School of Administration and Law, Dalian University of Technology, 2# Linggong Road, Dalian, 116085 (China)

ABSTRACT

Our objective in this paper is to explore how to measure technology convergence worldwide, by employing technology co-classification analysis, co-DC analysis, selecting patent publications from all the 20 sections in the world patent database Derwent Innovations Index, DII. We propose 3 different-level indicators and related methods on measuring technology convergence: Inner-section converging index (INSCI), Inter-section converging index (ITSCI) and Inter-category converging index (ITCCI), and employ these 3 indicators to conduct the empirical studies. Our findings not only can help us understand the status of technology convergence, but also can be used to facilitate our decision making.

INTRODUCTION

This paper aims at exploring how to measure technology convergence worldwide, trying to detect the state of art of technology convergence globally, by employing technology co-classification analysis, that is, co-Derwent Class Code analysis, i. e., co-DC analysis, a type of technology co-classification analyses, selecting patent publications from all the 20 sections in terms of Derwent Class Code System in the world patent database named Derwent Innovations Index, DII. This study will help us understand the current situation of the development of technology convergence, and afford somewhat objective results for our related decision-making.

Studies on measuring technology convergence appear to be pretty insufficient. The issue of measuring technology convergence has been attracting scientometricians' focus naturally recent years with more and more attention paid to the development of technology convergence over time.

DATA and TYPES of TECHNOLOGY CONVERGENCE in THIS PAPER

Data sources and database structure

Data in this study is from the most comprehensive world patent database, *Derwent Innovation Index, DII*. Derwent Class Code system, DC, is employed to finish data searching and analysing work. The structure of DC system in DII includes (Figure. 1): technology category (Abbr. as Category), technology section (Abbr. as Section), technology field (Abbr. as Field).

¹ This work was supported by the Project of National Natural Science Foundation of China (NSFC) under Grant 71473028; Key Project of the National Social Science Fund under Grant No. 13AZD016 and Project of the National Social Science Fund under Grant No. 14BTQ030.

The data we searched in this study covers all the three technology categories, i.e., Chemical, Engineering, Electric & Electronics; all the 20 technology sections, with DC system in DII. The sample data we analysed in this study includes all the patent publications in each of the 20 sections, or the top limit searching results of 100,000, published in 2014. The whole data sample covers 1,405,124 patent publications. Our searching time is during January 1-16, 2015.

Types of technology convergence in this study

Three types of technology convergence will be analysed and compared from micro-level, meso-level to macro-level. That is, Inner-section convergence, Inter-section convergence and Inter-category convergence, respectively (Figure. 2).

Figure 1: The structure of Derwent Class Code System in *Derwent Innovations Index, DII*.

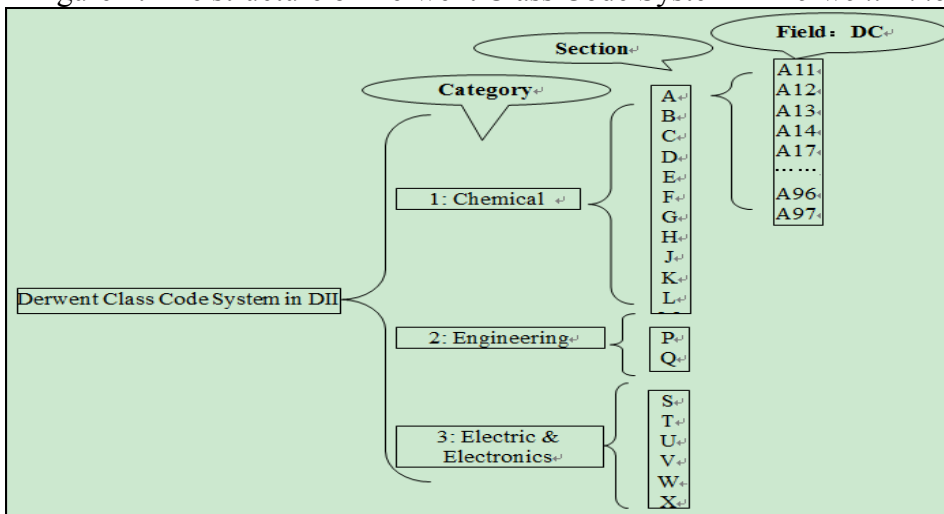
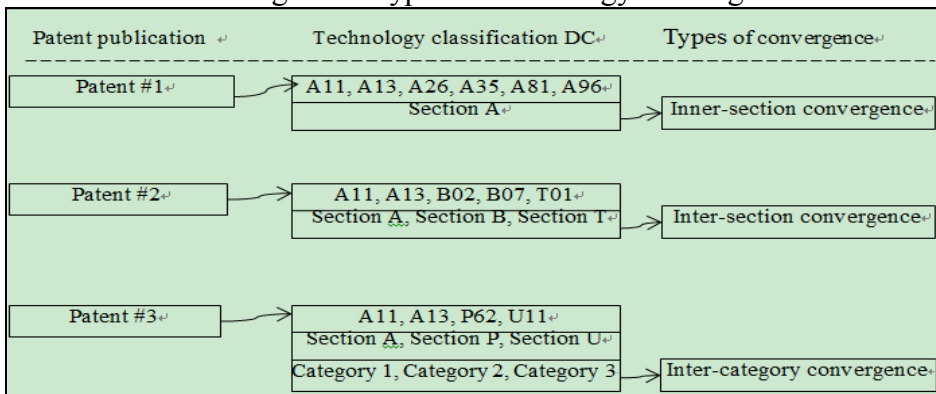


Figure 2: Types of technology convergence in this study.



METHODS and INDICATORS

General methods

Technology co-classification analysis, i. e., patent co-DC analysis in this study, is employed as the general method in the paper. In DII, each patent record is assigned with one DC or more DCs, such as B04, X21, P72 and L01, et al., for instance, if 10 DCs are assigned to a specific patent publication, then it indicates that this patent technology is a converging technology which integrates 10 DC technology fields; there exists a co-DC relationship among the 10 DCs, so technology co-DC analysis can be conducted in this investigation.

Indicator 1: Inner-Section Converging Index (INSCI)

1) Step 1: Basic statistics and computation

After searching work, and basic statistical work of DC distribution of each technology section from A to X has been done, and the total number of DCs to a specific section is computed and filled in the corresponding cell in the second line in Table. 1, such as ADCs represents the number of total DCs beginning with A in section A in terms of the selected sample data.

2) Step 2: Indicator of “inner-section converging index” in each section

According to the definition of “inner-section convergence” explained in the part of 2.2 “types of technology convergence in this study”, indicator 1 of “inner-section converging index”, abbr. as INSCI, can be calculated with the following Formula 1.1).

Table 1. Table captions should be centred and placed above the table.

Section	A	B	...	W	X
DCs	ADCs	BDCs	...	WDCs	XDCs
INSCI	INSCIA	INSCIB	...	INSCIW	INSCIX

$$INSCI(i) = (iDCs - Ri) / Ri \quad 1.1)$$

Where $INSCI(i)$ represents the indicator of “inner-section converging index” of technology section i ; i refers to technology section from A to X; $iDCs$ represents the number of total DCs beginning with i in section i in terms of the selected sample data; Ri denotes the number of the selected sample data in each section from A to X.

INDICATOR 2: INTER-SECTION CONVERGING INDEX (ITSCI)

1) Step 1: Basic statistics and computation

After searching work, and basic statistical work of DC distribution of each technology section from A to X has been done, count of DCs (denoted by a-x) distribution in other 19 sections has been computed and filled in the second column for each section from A to X (Tab. 3). AB, AC,, AW, AX,, A(i) (Formula 2.1), representing inter-section converging index, that is, index of section A converging section B, index of section A converging section C,, respectively. Take section A for example as following formula (Formula 2.1).

$$A(i) = iDCs / 100000 \quad 2.1)$$

Where i denotes section from B to X; $A(i)$ represents the index of section A converges other sections, such as AB denotes the index of section A converges section B; $iDCs$ represent the number of DCs from B to X.

2) Step 2: Indicator of “inter-section converging index” in each section

According to the definition of “inter-section convergence” illustrated in the part of 2.2 “types of technology convergence in this study”, taking Section A for example, indicator 2 of “inter-section converging index”, abbr. as ITSCI, and it can be calculated with the following Formula 2.2).

$$AITSCI(A) = ((AB) + \dots + (AX)) / 19 \quad 2.2)$$

Where $AITSCI(A)$ represents “average inter-section converging index” for section A: the first A denotes average, the second A denotes Section A. AB denotes the index of section A converges section B, \dots , AX denotes the index of section A converges section X, et al.. The number of 19 denotes other 19 sections excluding section A. With the same method, the values of “average inter-section converging index” in each section can be obtained and compared.

Indicator 3: Inter-Category Converging Index (ITCCI)

There are three categories in Derwent Class Code System in DII as shown in Figure. 1: Category 1, Chemical; Category 2, Engineering; Category 3, Electric & Electronics. So there exist total 6 permutation in terms of measuring inter-category converging index, i. e., $ITCCI (3 \rightarrow 1)$, representing “inter-category converging index” of Category 3 converges Category 1; on the contrary, $ITCCI (1 \rightarrow 3)$, representing “inter-category converging index” of Category 1 converges Category 3, and so on. Here, we take $ITCCI (3 \rightarrow 1)$ as an example to illustrate the steps, formulas and indicators concerning the issue of “inter-category converging index” measurement.

1) Step 1: Basic statistics and computation

Total 6 Sections are in Category 3, that is, S, T, U, V, W and X; total 12 sections are in Category 1, i. e., A, B, C, D, E, F, G, H, J, K, L and M. First, the count and converging index of each section in Category 3 converges each section in Category 1 are counted and computed, such as SA denotes the converging index of Section S converges Section A, by such analogy.

2) Step 2: Indicator of “inter-category converging index”

According to the definition of “inter-category convergence” explained in the part of 2.2 “types of technology convergence in this study”, indicator 3 of “inter-category converging index”, abbr. as $AITCCI (3 \rightarrow 1)$, can be calculated according to the following procedures.

First, the value of converging index of Section S converges each section from A to M in Category 1 (abbr. as S1) can be summed up according to Formula 2.3) as follows.

$$S1=SA+SB+SC+SD+SE+SF+SG+SH+SJ+SK+SL+SM \quad (\text{omitted as } S1= SA+\dots+ SM \text{ in Tab.4}) \quad 2.3)$$

Where S1 denotes the sum of converging index of Section S converging each section from A to M in Category 1, that is, the converging index of Section S converges Category 1. Referencing the same method, T1 to X1 can be calculated, respectively.

Then, “average inter-category converging index” of Category 3 converges Category 1 [abbr. as $AITCCI (3 \rightarrow 1)$], can be figured out according to the Formula 2.4) as follow.

$$AITCCI (3 \rightarrow 1)=(S1+T1+U1+V1+W1+X1)/6 \quad 2.4)$$

Where $AITCCI (3 \rightarrow 1)$ represents average inter-category converging index of Category 3 converges Category 1; $S1+T1+U1+V1+W1+X1$, denotes the sum of converging index of each

section in Category 3 converges each section in Category 1; 6 denotes total 6 sections in Category 3.

EMPIRICAL ANALYSES and RESULTS

Result of indicator 1: Inner-Section Converging Index

According to the method and indicator illustrated in the part of 3.1, the values of indicator 1: inner-section converging index in each section from A to X have been figured out, and then these values are sorted by descending order, and drawn in Figure 3.

Figure 3: Indicator 1-Inner-section converging index by descending order.

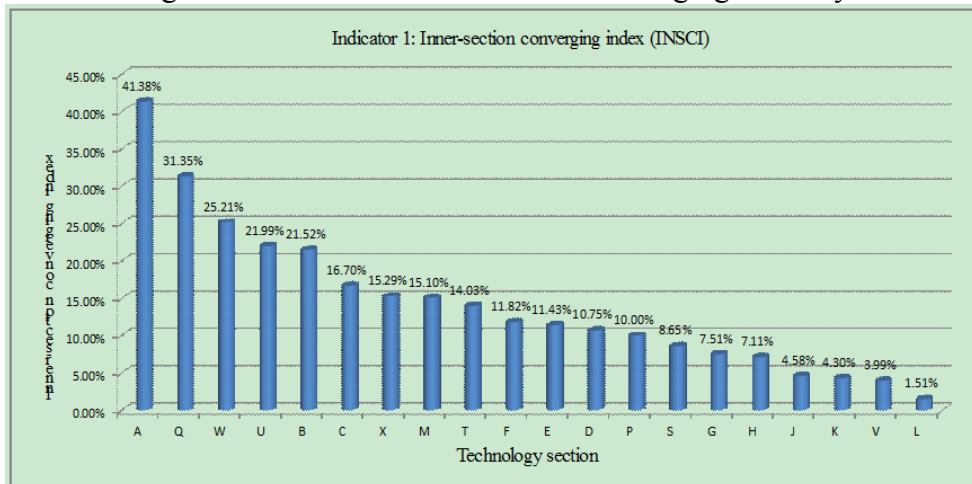


Figure 3 shows that the value of “Inner-section converging index” (INSCI) in section A is the highest, 41.38%, much higher than that of the other sections’, indicating that knowledge flow and convergence are much frequently and actively within section A. The value of INSCI in section Q is the second highest, 31.35%. The value of INSCI in another 3 sections, section W, section U and section B, is over 20%. Section A and section B in top 5 are members of Category 1, Chemical; section Q in top 5 is a member of Category 2, Engineering; section W and section U in top 5 belong to Category 3, Electrical & Electronic. There are 7 sections with value of INSCI lower than 10%, that is, section S, section G, section H, section J, section K, section V and section L, 5 of them being members of Category 1, Chemical, 2 of them being members of Category 3, Electrical & Electronic. The last one with the lowest value of INSCI (1.51%) is section L.

Result of Indicator 2: Inter-Section Converging Index

According to the method and indicator illustrated in the section of 3.1- “General method” and 3.3- “Indicator 2: inter-section converging index (ITSCI)”, after statistical and computing work, the average values of indicator 2: inter-section converging index in each section from A to X have been worked out. Then these values are sorted by descending order, and drawn in Figure 4.

Figure 4: Indicator 2-Inter-section converging index by descending order.

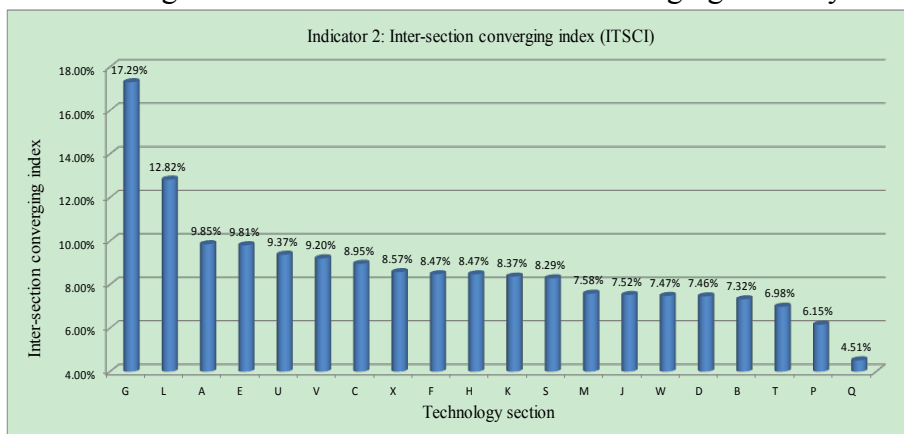


Figure 4 shows that Section G, has the highest value of average inter-section converging index, ITSCI, 17.29%, being far ahead than that of in other 19 Sections, reflecting much more knowledge from other Sections has been integrated or absorbed into Section G, or the achievements of Section G have been widely applied in other Sections. Section L, has the second highest value of average ITSCI, 12.82%. Top 4 Sections including Section G, Section L, Section A and Section E, are all members of Category 1, Chemical, reflecting some of Sections among Category Chemical have an active knowledge flow and communication with other technology Sections. The last 2 Sections are Section P, and Section Q, both of them being with much lower average values of ITSCI, 6.15% and 4.51%, respectively, reflecting low level knowledge flow and integration for the 2 Sections of Engineering technology with other technology Sections.

Result of Indicator 3: Inter-Category Converging Index

For there exist total 6 permutations in terms of measuring inter-category converging index, first, we take Category 1 converging Category 3, that is, ITCCI (3 → 1) as an example. According to the method and indicator illustrated in the section of 3.1- “General method” and 3.4- “Indicator 3: Inter-Category Converging index (ITCCI)”, after statistical and computing work, according to the formula 2.4), the average value of *ITCCI (3 → 1)* has been worked out.

By employing similar method, other 5 average *ITCCIs*, *AITCCI (1 → 3)*, *AITCCI (3 → 1)*, *AITCCI (1 → 2)*, *AITCCI (2 → 1)*, *AITCCI (2 → 3)*, *AITCCI (3 → 2)*, have been figured out as shown in Table 2.

Table 2. Results of indicator 3: Inter-Category Converging Index (*ITCCI*).

Inter-Category	AITCCI	Average
Category 1 → Category 2	AITCCI (1 → 2): 37.61%	39.62%
Category 2 → Category 1	AITCCI (2 → 1): 34.85%	
Category 2 → Category 3	AITCCI (2 → 3): 53.90%	
Category 3 → Category 2	AITCCI (3 → 2): 37.47%	
Category 3 → Category 1	AITCCI (3 → 1): 32.35%	
Category 1 → Category 3	AITCCI (1 → 3): 41.55%	

Table 2 shows that the average value of Inter-Category Converging Index is 39.62% between any 2 Categories. The activity of Category 2 converging Category 3 appears to be most dynamic, the converging index of *AITCCI* (2 → 3) is up to 53.90%, reflecting Category Engineering absorbs and integrates plenty of knowledge from Category Electrical & Electronic, or a big amount of achievements in Category Engineering are applied in Category Electrical & Electronic. The converging index of *AITCCI* (1 → 3), Category Chemical converging Category Electrical & Electronic, is the second highest, 41.55%, indicating a comparable converging activity taken place between these 2 Categories. The last one with the lowest *AITCCI* is *AITCCI* (3 → 1), the index of Category Electrical & Electronic converging Category Chemical, 32.35%.

CONCLUSIONS and DISCUSSIONS

Conclusions

We proposed three different-level indicators: 1) Inner-Section Converging Index (INSCI), to measure technology convergence taken place among technology fields within a same technology section; 2) Inter-Section Converging Index (ITSCI), employed to measure technology convergence occurred among different technology sections; 3) Inter-Category Converging Index (ITCCI), used to measure technology convergence taken place among different technology categories.

Empirical studies show that, for indicator “Inner-Section Converging Index” (INSCI), section A has the highest value, indicating that knowledge flow and convergence are pretty frequently and actively within section A. The value in Section Q is the second highest. There are 7 sections with value of INSCI lower than 10%, that is, section S, section G, section H, section J, section K, section V and section L. From technology Category level, “average value of INSCI” in Category 2, Engineering, is the highest, 20.67%; INSCI in Category 1, Chemical, is the lowest.

When it comes to indicator 2), “Inter-Section Converging Index” (ITSCI), technology Section G has the highest value, leaving other 19 sections far behind. Section L, has the second highest value. The last 2 Sections are Section P and Section Q.

As for indicator 3), Inter-Category Converging Index (ITCCI), the activity of Category 2 converging Category 3 appears to be most dynamic, the converging index is up to 53.90%. While the index of Category Electrical & Electronic converging Category Chemical is the lowest, 32.35%.

Discussions

The innovation of this paper lies in our proposing 3-different-level indicators on measuring technology convergence globally, and employing them to finish our empirical analyses. This study helps us understand the current situation of the development of technology convergence, and afford somewhat objective results for our related decision-making.

The theoretical significance of this study is that our proposing 3 different-level indicators and related methods on measuring technology convergence: Inner-Section Converging Index (INSCI), Inter-Section Converging Index (ITSCI), and Inter-Category Converging Index

(ITCCI). These indicators can be employed to measure technology convergence from micro-level, meso-level and macro-level, respectively, and the measuring results can help us comprehend the developmental stage of technology convergence worldwide. Our findings not only can help us understand the status of technology convergence, but also can be used to facilitate our decision making.

REFERENCES

- Altshuller, H. (2001). *And Suddenly the Inventor Appeared: TRIZ, the Theory of Inventive Problem Solving*. Worcester, Massachusetts, *Technical Innovation Center, INC*.
- Ash, J. (1997). Organizational factors that influence information technology diffusion in academic health sciences centers. *Journal of the American Medical Informatics Association*, 4(2): 102-111.
- Ausloos, M. (2015). Coherent measures of the impact of co-authors in peer review journals and in proceedings publications. *Physica a-Statistical Mechanics and Its Applications*, 438: 568-578.
- Berenson, C. (1968). R AND D - Marketing interface - general analogue model for technology diffusion. *Journal of Marketing*, 32(2): 8-15.
- Blank, J. L. T. and V. G. Valdmanis (2015). Technology diffusion in hospitals: a log odds random effects regression model. *International Journal of Health Planning and Management*, 30(3): 246-259.
- Callon, M., J. P. Courtial, W. A. Turner and S. Bauin (1983). From translations to problematic networks - an introduction to co-word analysis. *Social Science Information Sur Les Sciences Sociales*, 22(2): 191-235.
- Curran, C. S. and J. Leker (2011). Patent indicators for monitoring convergence - examples from NFF and ICT. *Technological Forecasting and Social Change*, 78(2): 256-273.
- Geum, Y., C. Kim, S. Lee and M. S. Kim (2012). Technological Convergence of IT and BT: Evidence from Patent Analysis. *Etri Journal*, 34(3): 439-449.
- Luan, C., Z. Liu and X. Wang (2013). Divergence and convergence: technology-relatedness evolution in solar energy industry. *Scientometrics*, 97(2): 461-475.
- Magalhaes, M. and C. Hellstrom (2013). Technology diffusion and its effects on social inequalities. *Journal of Macroeconomics*, 37: 299-313.
- Marinakos, Y. D. (2012). Forecasting technology diffusion with the Richards model. *Technological Forecasting and Social Change*, 79(1): 172-179.
- Muddiman, D. (2003). Red information scientist: the information career of J.D. Bernal. *Journal of Documentation*, 59(4): 387-409.

Nelson, R. R., E. Denison, K. Sato and E. S. Phelps (1966). Investment in humans, technological diffusion, and economic growth. *American Economic Review*, 56(2): 69-82.

Ravikumar, S., A. Agrahari and S. N. Singh (2015). Mapping the intellectual structure of scientometrics: a co-word analysis of the journal *Scientometrics* (2005-2010). *Scientometrics*, 102(1): 929-955.

APPENDIX

Table 3. All 20 technology classifications (abbr. as DC).

(A - M)	Chemical Sections
A	Polymers and Plastics
B	Pharmaceuticals
C	Agricultural Chemicals
D	Food, Detergents, Water Treatment and Biotechnology
E	General Chemicals
F	Textiles and Paper Making
G	Printing, Coating, and Photographic
H	Petroleum
J	Chemical Engineering
K	Nucleonics, Explosives and Protection
L	Refractories, Ceramics, Cement and Electro(in) Organics
M	Metallurgy
(P - Q)	Engineering Sections
P	General
Q	Mechanical
(S - X)	Electrical & Electronic Sections
S	Instrumentation, Measuring and Testing
T	Computing and Control
U	Semiconductors and Electronic Circuitry
V	Electronic Components
W	Communications
X	Electric Power Engineering



Development on the periphery: monitoring science, technology and innovation for sustainable development among Pacific Island Countries¹

Ranasinghe M.W. Amaradasa* and Tim Turpin**

*Ranasinghe M.W. Amaradasa,
University of Fiji, Private mail bag, Saweni, Lautoka, Fiji.
wasanthasuriya@gmail.com

**Tim Turpin,
Adjunct Professor, School of Business Western Sydney University, Locked Bag 1797, Penrith NSW 2751
Australia.
t.turpin@uws.edu.au

ABSTRACT

This paper reviews the status of science, technology and innovation indicators in Fiji and other Pacific Island countries. Data are drawn from interviews with senior officials in Fiji, regional policy documents, and data held at the University of the South Pacific. The limited data available is mostly held in separate national agencies with little national or regional collaboration. The paper argues that the paucity of S&T data available for policy making or analysis is symptomatic of the nature of development in the region and the inappropriateness of indicators designed primarily for industrialised economies. It concludes with an observation that the drive toward sustainable development is steering a regional move toward development of an S,T&I indicator hub located across one or more Pacific Island countries.

¹ The authors wish to acknowledge a grant received from the University of Fiji to undertake fieldwork for this project: *Assessment of S&T capabilities in Fiji*

Introduction

The Pacific Island Countries (PICs) consist of 21 small island nations. Economically and in terms of science, technology and innovation (S,T&I) capability they are all very much on the periphery of economic and scientific development. Palau, has the highest GDP per capita of the PICs recording \$15, 100 in 2013 while Fiji recorded only \$7,800.² Many of the other member countries have rates, well below \$3,000. In contrast they are surrounded by the US and Canada to the east and Australia and New Zealand to the south and west with GDP rates for the same year varying from \$35,000 - \$53, 000. To the north Japan recorded a rate of \$36,000 while Singapore to the west recorded over \$78,000 (UNESCO, 2015).

The PIC economies are mostly dependent on natural resources, with very little manufacturing and no heavy industry (Turpin *et al.* 2015). They have small domestic markets and trade balances are heavily skewed toward imports. Although there have been recent collective approaches to science and technology policy development, there are, as yet, no comprehensive systematic data available to indicate the technical capacity of personnel or levels of expenditure on R & D and areas of research investment and innovation. The only systematic data included in the UNESCO Science Report (2015) were scientific publications data.

However, there are some data maintained by some national agencies. The present paper reports on data collected from senior policy personnel in government agencies, mostly in Fiji, concerning the status of science, technology and innovation across the region. The data suggest financial and personnel investments in science are low by international standards. Moreover, the institutions typically involved in the distribution of knowledge and innovative processes include many that are quite different from those that dominate the more industrialised economies. Our argument here is that the paucity of general S&T data is to some extent symptomatic of the nature of development across the region. However, it also reflects the limitations of S&T indicators, designed essentially for industrialised economies, to monitor or evaluate S,T &I development across these Island States.

Observers have drawn attention to development challenges facing PICs, such as distance from potential markets, a high dependence on foreign aid, limited or unstable natural resource, fragile eco-systems, small domestic markets, a reluctance of professionals to move into the region, and, in particular, cultural and institutional resistance to change (Tibben and Tielu, 2007; Hughes 2003 and Higgens, 1994.) Yet, somewhat paradoxically PICs are geographically at the centre of the economically and technologically dynamic Asia Pacific Economic community (APEC) that includes many of the wealthiest and technologically advanced countries in the world.

In this paper we seek to explore this contradiction and the current status of S,T&I indicators in the Pacific periphery. We sketch out some of the emerging science, technology and innovation (S,T&I) policy issues across PICs and the indicators likely to be most useful for monitoring sustainable development into the future.³

² GDP per capita, current PPP, 2013 (UNESCO, 2015)

³ Interviews were held with officials representing: the Ministry of Strategic Planning; the Bureau of Statistics; the Ministry of Education and Department of Technical Education; the Ministry of Trade; the Ministry of

Challenges for S,T&I indicator development among PICs

A large informal economy

Much of the production across the PICs takes place in the informal economy. Small scale and semi-subsistence farming dominates agriculture. In 2003 only 20 per cent of the population in Fiji were employed in the formal economy while in Samoa less than 15 per cent were in the formal economy (Tibben and Tielu, 2007). In PNG around 85 per cent of the labour force are engaged in agriculture and over 80 per cent so employed in Vanuatu (ADB, 2003). As a consequence, S&T data are difficult to collect and what can, is unrepresentative of what actually takes place in productive sectors.

With such a high proportion of the workforce in subsistence or self-employed production links between local knowledge and scientific or technological developments are weak. Such links are compounded by a range of factors including local social and political processes and traditional land tenure processes. A recent sustainable energy action plan focus group report (SEAP, 2012) noted that the wider cultural and social context heavily influences S&T. Further, limited freedom of expression and in some cases religious settings discourage research in certain areas. Figure 1 (following Tibben and Tielu (2007) presents a stylized representation of institutional structures engaged, one way or other, in the distribution of knowledge and innovation processes in Fiji. The schema depicted here is somewhat different from the representations of national, innovation systems presented elsewhere (see for example, Galli and Teubal, 1997 or Barnard et al, 2009). Among PICs local and Indigenous organisations play an important role as a clearing house for knowledge but they can be in conflict with demands from other institutional processes (Tibben and Tielu, 2007).

Yet, improving knowledge flows between the practices and expectations of farmers, fishers, foresters, related producers and the nascent scientific and research community remains a critical task. The system depicted in Figure 1 is well into a transition process. For example the University of the South Pacific's School of Agriculture supports an institute for agricultural extension and research and the Secretariat of the Pacific Community (SPC) has a similar extension and agricultural development program. Activities in this context lie at the heart of local innovation processes but there is little in the way of indicators to monitor or evaluate progress. Relationships between local communities and the mining activities in PNG reflect similar challenges. Such issues become less important in urban production settings.

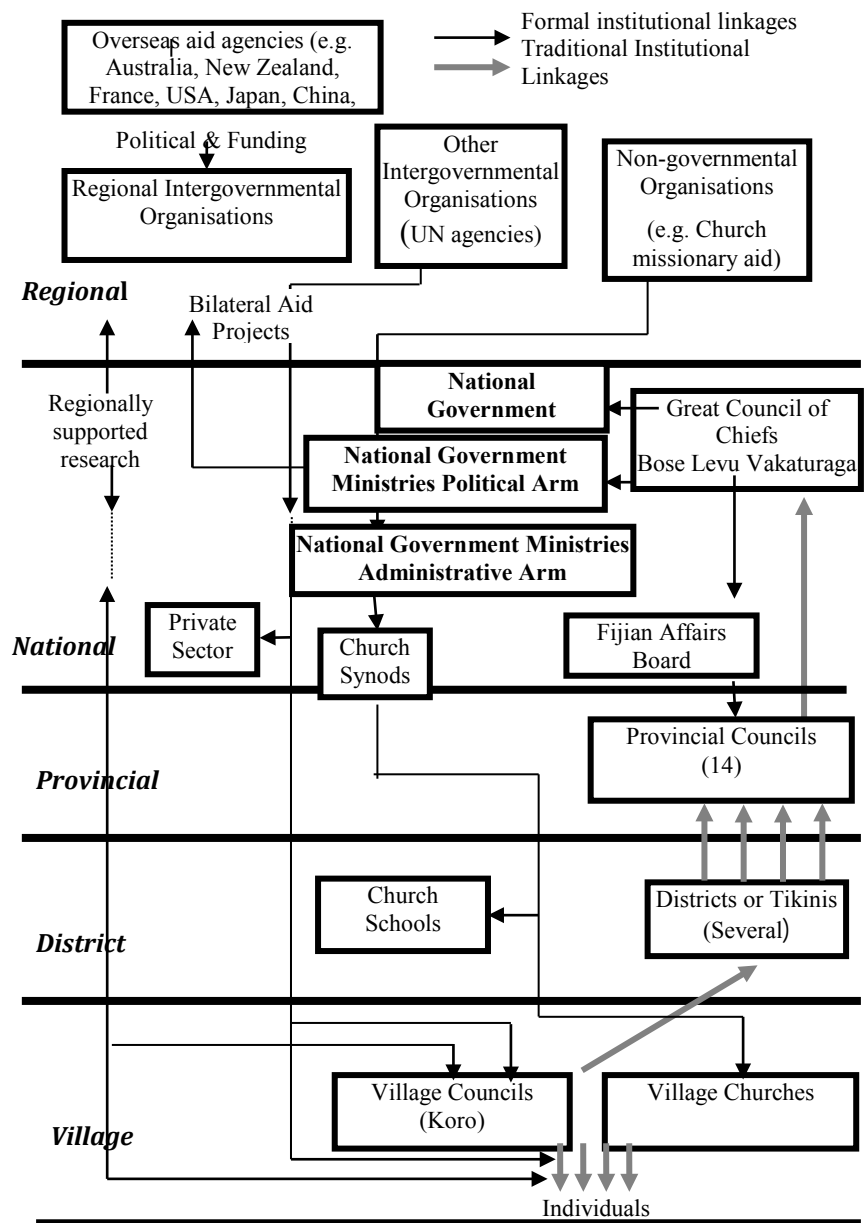
Limited policy frameworks

While there is considerable diversity across PICs they share many common development challenges. Recently regional issues have led to some collective policy action and progress toward monitoring regional capability. These have focused mainly on climate change, natural disaster from hurricane or tsunami, energy and telecommunications, fisheries, agriculture, telecommunication and environmental sustainability.

Health; and the Ministry of Agriculture in Fiji. We also draw on our analysis of recent regional policy documentation.

Examples of collective policy action include the work of the Secretariat of the Pacific Community (SPC) on climate change, fisheries and agriculture; The Pacific Forum Secretariat for transport and telecommunications; and the Secretariat of the Pacific Region Environmental Programme. None of these, however, deal directly with S&T issues and there is little demand or direction for related data collection.

Figure 1: Institutional linkages among Fiji's institutions



Source: Tibben and Tielu, 2007 p. 391.

More recently the Pacific-Europe Network for Science Technology and Innovation (PACE-Net Plus), funded by the European Commission, seeks to reinforce the dialogue between the Pacific Region and Europe in ST&I. Key areas of focus identified so far include: health; food

security, sustainable agriculture; marine and maritime research; climate action and resource efficiency (Turpin *et al.* 2015). This new initiative underscores the importance of collective action among PICs and their need for foreign funds to build international science networks.

Sources of S&T funding

The majority of S&T investment across PICs is derived from foreign aid. Clearly such investments serve to forge and develop institutional S&T platforms and have been instrumental in developing capacity in institutions such as the University of the South Pacific and its regional campuses. However, the institutional processes through which S&T finances are directed have come under criticism (Kelsey, 2004). While such collaboration may potentially benefit PIC researchers and those in other parts of the globe, there is considerable debate about the extent to which this outcome benefits national development within the region. Table 1 illustrates an historical connection between the PIC countries and the location of their collaborating partners. A policy question for higher education research is whether these are likely to continue to be highest priority for future domestic research support. For example, would more collaboration with countries in Southeast Asia with strong agricultural research programmes, such as the Philippines, deliver greater practical benefit? Answers to such questions require a rather different approach to bibliometric analyses than is usually presented in global data.

Table 1: Top three countries for international co-authorship with Pacific Island Countries

PIC	International co-authors (%)	Top three o-authorship countries (%)		
Fiji	80.0	Australia (33.9)	USA (16.2)	New Zealand (13.6)
French Polynesia	91.5	France (59.6)	USA (22.9)	Australia (11.9)
New Caledonia	87.2	France (51.1)	USA (16.4)	Australia (16.2)
Papua New Guinea	84.8	Australia (45.7)	USA (28.0)	United Kingdom (14.5)

Source: SPU with data derived from Thompson's web of science.

S&T indicators for innovation and national development: the case of Fiji.

Building on the regional natural resource base

The information gathered for the present project reveals some of the S,T&I indicator challenges facing the peripheral PICs. Sustainable development strategies seek to move the economy beyond primary production toward value added production and marketing. However, this requires S&T input in order meet export standards.

The 'Fiji 2020 Agriculture Sector Policy Agenda' (MOAF, 2014) highlights the need for a high priority on food security by modernizing agriculture and introducing integrated systems to include improved delivery of technical support. This includes curriculum development, research, and development of innovative agricultural business models to promote commercial agriculture.

Forestry, the fifth highest contributor in Fiji to national export earnings and a major export product for PNG uses low and semi intensive technological inputs. As a consequence product ranges are limited to sawn timber, veneer, plywood, block board, molding, poles and post, and wood chips. Only a small amount of finished products are exported; most contribute to the local hotel industry. Lack of automated machinery coupled with inadequate technically skilled local workers to operate automated machinery and limited designing capabilities inhibit more intensive value-added exploitation.

Most of the commercial energy needs throughout the region are currently met from fossil fuel. However, Fiji has significant shares (60%) of renewable energy in the total electricity supply. Many PICs have identified the role of renewable energy sources towards social development and set ambitious goals with plans to have renewable energy generating a much larger share of the electricity supply. Fiji aims to be 90% reliant on renewable energy while Tonga, Nauru and the Cook Islands aim to be at least 50 per cent by the end of 2015 (SPC, 2013). S&T capabilities will be an important factor for achieving this and monitoring data on progress is likely to be of common interest across all PICs. While some data for such policy development, monitoring or evaluating are held in individual agencies it does not appear to be comprehensively available across government, nor across the PICs. Data sharing nationally or internationally appears limited.

Toward S&T indicators for PICs

A recent education policy initiative in Fiji was the establishment of the Higher Education Commission (FHEC) in 2010 which is the regulatory body charge of tertiary education in the country. FHEC has embarked on a registration and accreditation processes for tertiary level education providers to improve the quality of higher education. Further, since 2014, the FHEC allocates research grants to academic researchers. The Ministry of Education holds data on enrollments in science.

The Pacific Island Countries have experienced positive annual growth in scientific output with an average growth of 17 per cent per year. A database, maintained at USP, shows a significant growth from 43 publications in 2000 to 282 in 2013 (an average growth rate of 9 per cent per year). However, the number of articles and book chapters published with an affiliation to an author in a PIC has not grown substantially. Scientific output is dominated by the medical sciences, agriculture and biological sciences. There is a strong concentration on immunology/genetic/biotechnology /microbiology in PNG. Citation rates remain favorably comparable with other parts of the world (UNESCO, 2015).

Interpretation of collaborative data has led some analysts to argue there is very little in the way of direct return to Fiji. This argument has led to a new set of guidelines in the health sector to extend training and access to new technology and to ensure that domestic research funding is consistent with building local research capacity.

Some regional S&T targets have been identified for collective action among PICs. PACE- Net (PACE-Net 2013) for example, has provided recommendations for a Strategic Research Plan for Research, Innovation and Development in the Pacific. The plan emphasizes seven key areas:

1. Agriculture & Forestry;
2. Health;
3. Fisheries and aquaculture;
4. Biodiversity & Ecosystem management;
5. Fresh water;
6. Natural Hazards; and
7. Energy. (PACE-Net recommendations, 2012).

A Pacific Island Research Network was established in 2012 to discuss the formation of a network of universities from the Pacific Island countries and territories to support intra- and inter-regional knowledge creation and sharing, and to prepare succinct recommendations for the development of a regional ST&I Policy framework.

Fiji has recently developed a Green Growth Framework which in Aug 2014 emphasised sustainable development with a focus on technology and innovation. A number of implications for S&T data collection for monitoring progress follow from this initiative. For example, technological subsidies to firm; introduction of tariffs on non-green technologies; incentives to attract FDI for developing environmentally sustainable technologies as proposed in the Fijian Green Growth Framework clearly require data for monitoring and evaluation.

To a considerable extent Fiji has become a PIC hub for education and training but as yet it has not become a hub for data collection, data sharing or collective analysis. With the development of a regional university networks and the capacity of USP to monitor both bibliometric and research project data there is the potential for concomitant regional data analysis.

Conclusions

The summary above shows that Pacific Island Countries have much in common. All are marginalised or peripheral to the extensive scientific and technological progress taking place in the countries surrounding them. S,T&I national policy initiatives are evolving in some PICs and there is evidence of emerging regional policy developments. These require data for monitoring and evaluation. However, as the PIC case shows, data demands across this region are different from those in the industrialised core. For example, the institutional actors in the schematic representation of Fijian institutions are different from the typical industrial national innovation system. Data concerning the process through which these institution are evolving and the future role they can play in development strategies is important for both policy planning and evaluation.

While inward flows of professional and skilled personnel are important for the industrialised countries that surround them, for PICs the outward flow is of far more strategic importance. Further, data on outward personnel flows could be used to design policies for more strategically oriented international scientific collaboration.

At present the economic and social structures in place provide little demand for the costly process of data collection and analysis. However, as regional based policy initiatives are evolving there is likely to be increasing demand for collective approaches to data collection.

While some globally comparative data, such as R&D investment and numbers of researchers and technicians can serve some useful comparative processes, it is the more locally specific data that offer greater policy returns from investment. For example, in the agriculture and resource based sectors data on the work of extension officers is likely to return more economically useful information than innovation data derived from firms. For the latter, data concerning skills demands for value-added production, albeit at a quite basic level, are likely to be informative for education and training policy and technological acquisition.

In short, collective S,T&I policy action among PICs has tended to move faster than national or collective approaches to data collection and analysis. It is likely that for some PICs there will, at their national level, continue to be little demand for S,T&I indicators. The cost is too high and the content of limited relevance. For them, the data emphasis is likely to remain focused on the broader concepts of poverty reduction and sustainable development generally for quite some time. In the longer term, however, cooperation between PICs may well lead to the formation of a regional data monitoring hub involving one or more of the larger PIC economies. The current evidence suggests this process is already underway.

References

- ADB (2003) *Key Indicators 2003 Education for Global Participation*, Manila, Asian Development Bank.
- Barnard, Helen, Tracy Bromfield and John Cantwell (2009) 'The role of indigenous firms in innovation systems in developing countries: the development implications of national champion firms' response to underdeveloped national innovation systems' in Bengt-Akile Lundvall, K.J., Joseph, Christime Chaminade and Jan Vang (eds.) *Handbook of Innovation Systems and Developing Countries*, Edward Elgar, Cheltenham.
- Galli, Riccardo and Morris Teubal (1997) 'Paradigmatic Shifts in National Innovation Systems' in Charles Edquist (ed.) *Systems of Innovation: Technologies, Institutions and Organisations*, Pinter, Ottawa.
- Higgins, B. (1994) 'Technology and Economic Development', in A. Marjoram (ed.) *Island Technology: Technology for Development in the South Pacific*, London, Intermediate Technology Publications.
- Hughes, H. (2003) 'Aid has failed the Pacific', Centre for Independent Studies, St Leonards, NSW, Australia.
- Kelsey J. (2004) 'Big brothers behaving badly; the implications for the Pacific Islands of the Pacific Agreement on Closer Economic Relations (PACER) Pacific Network on Globalisation, Suva, Fiji

PACE-NET, (2013) 'Recommendations for a strategic plan on research, innovation and development in the pacific', available at <http://pacenet.eu/news/pacenet-outcomes-2013> accessed on 17th Oct 2014.

PACE-NET (2014), Pacific Islands Universities Research Network (PIURN) communiqué , available at <http://pacenet.eu/news/pacenet-outcomes-2013> accessed on 17th Oct 2014.

SPC (2013), SPC Climate change and disaster risk management support activities in Pacific Island Countries and territories, compiled by the Secretariat of the Pacific Community (SPC) ISBN: 978-982-00-0646-1. SPC, New Caledonia.

Tibben, William and Apela Tielu (2007) 'Innovation Strategy for the Pacific Islands in the new millennium: mixing science with tradition', in *Science Technology Policy and the Diffusion of Knowledge*, (Eds.) Tim Turpin and V.V. Krishna, Edward Elgar, Cheltenham.

Turpin, Tim, Jing, J. Zhang, Bessie M. Burgoss and Wasantha Amaradasa (2015) 'Southeast Asia and Oceania', in *UNESCO Science Report: Towards 2030*, (ed.) Susan Schneegans, UNESCO, Paris.

UNESCO (2015) *UNESCO Science Report: Towards, 2030* (ed.) Susan Schneegans, UNESCO, Paris.

CHAPTER 14

Networks and Text Analysis



Breakout discoveries in science: what do they have in common?

Jos J. Winnink^{*}, Robert J.W. Tijssen^{*} and Anthony F.J. van Raan^{*}

{winninkj, tijssen, vanraan}@cwts.leidenuniv.nl
Leiden University, Centre for Science and Technology Studies (CWTS),
Wassenaarseweg 62A, 2333 AL Leiden (the Netherlands)

ABSTRACT

We applied a series of computerised search algorithms to identify potential breakthrough discoveries in science. These citation-based ‘early detection’ algorithms extract their information from high-quality bibliographic databases. With early detection we focus on a paper’s citation pattern during the first 24-36 month after publication. In this study we apply those algorithms to large sets of publications from 2007-2011 in the Web of Science Core Collection database, and generated a dataset consisting of publications that are considered high-impact papers. We call these papers ‘breakouts’. We present general bibliographic characteristics in an attempt to classify communalities and develop a general typology of our identified ‘breakout’ publications in worldwide science. Our first results reveal a significant overrepresentation of various kinds of cross-sectorial collaborative research. The first results also reveal that ‘breakout’ publications are produced by on average larger teams of researchers and that this size depends on the type of collaboration.

INTRODUCTION

As technological change is commonly considered the main determinant of economic growth since seminal work in the 1950s and 1960s (Smits, 2002) many governments focus on policy measures or policy programs to stimulate technological innovation (OECD, 1992). To develop public policies to stimulate technological developments timely knowledge of new and promising technologies is essential. Martin (1995) states that policy-makers and scientists are grappling with the problem of how to select the most promising research areas and emerging technologies. Therefore one of the topics in economic policy making is the identification and forecasting of emerging new technology fields. Bettencourt et al. (2009, p. 219) say ‘... a more quantitative “science of science” may allow society to reap the benefits of new discoveries sooner ...’. Gathering information of new scientific discoveries at early stage in a reliable way could make timely well-founded decisions with respect to science and technology stimulation programs, and R&D investments feasible.

It is well known that only a small number of scientific discoveries stand out, and have an impact leading to major structural changes in science. Such rare discoveries are often referred to as ‘breakthroughs’. Given the vast number of scholarly publications published each year an automated computerised selection system might be a preferable method to harvest databases with bibliographic data of scholarly publications and to search for high-impact publications. Such a generalized and transparent method should facilitate the early and unbiased detection of potentially important new directions in science and technology. An objective method, consisting of one or more algorithms, is relevant as human beings who carry out the

evaluation of new developments might be forced to follow a set of strict protocols. The role of these protocols is to prevent preconceptions that could influence this process of evaluation.

Searching for breakthrough-class discoveries in science using bibliographic information is done for instance by Redner (2005), Schneider and Costas (2015), and Ponomarev et al. (2014). During the last few years we developed and implemented automatic computer algorithms that are based on characteristic patterns found in the early-stage citation history of breakthrough papers (Winnink et al., 2016). We use these algorithms to harvest large set of bibliographic data searching for scholarly papers that stand out. Such papers are selected on the basis of the presence of characteristic patterns in a papers citation profile during the first 24-36 months after publication. Because of this early recognition and the relative short time period taken into account the algorithms are unable to unmask hypes, hoaxes¹, and frauds². We denote those automatically selected papers ‘breakout’ papers as expert opinions and time are needed to cast a judgement if these research findings represent a broadly accepted scientific breakthrough.

In this study we focus on the typology of the organisations (co) producing breakout papers. To do so we analyse scholarly papers on several dimensions. The factors we currently investigate are (1) the type of the organisations with which collaborating authors are affiliated, (2) cross border collaboration –local, domestic, and international-, and (3) the size of the research group, (4) the effect of collaboration in different science fields. and (5) the characteristics of papers where the breakout patterns not occurs at the moment of publication but at a later stage. In this paper we focus on the type of the organisation or the organisational cooperation that produces the paper, the science field the paper is in, and the size of the author team of the paper. We use the earliest point in time after publications a paper is identified as a ‘breakout’ paper by one of the algorithms as point of reference.

METHODOLOGY

We applied to the data in the CWTS-licenced in-house version of Thomson Reuters Web of Science database (WoS) the five computerised algorithms we developed and implemented (Winnink et al., 2016). All papers of WoS-type ‘article’ and ‘letter’ that were published between 2007 and 2011 were selected. The application of the algorithms resulted in a set of papers each classified as a breakout paper by at least one of the algorithms. In (Winnink et al., 2016) we show that the algorithms prefer to select those papers that after 20 odd years turn out to be high impact papers. In this preliminary study we applied the algorithms to citation profiles in the first 24-36 months after publication. This approach proofed to be sufficiently reliable, as we found that 92% of all papers that show a ‘breakout’ character in the first ten years after publication are identified.

¹ An example of a hoax is the claim for the existence of nuclear fusion at room temperature - ‘cold fusion’ (Fleischmann, M, and Pons, S., 1989).

² For instance fraud committed by the Korean researcher Hwang Woo-Suk who was considered one of the pioneering experts in the field of stem cell research until a publication by David Cyranoski (2004) uncovered Hwang’s fraudulent research.

TOWARDS A TYPOLOGY: SOME PRELIMINARY RESULTS**Effect of organisational collaboration across science fields on the production of breakout papers**

We found 38,949 breakout publications among 4,799,020 papers for the period 2007-2011. The results are differentiated along two dimensions (Table 1). The first dimension (Organisational category) contains the organisational-type of the affiliation of the authors and combinations thereof. The second dimension consists of the categories that form the highest level of the NOWT³-classification; we left out the categories ‘Language, Information and Communication’, and ‘Law, Arts and Humanities’ as for these fields we found less than 15 breakout papers. The overall share of breakout papers we find is 0.8%.

Table 1 Distribution of breakout papers across NOWT-categories (articles + letters, 2007-2011)*

	Medical and Life Sciences	Natural Life Sciences	Engineering Sciences	Social and Behavioural Sciences	and Multidisciplinary journals**
Total number of papers	2,362,512	2,066,802	508,093	360,443	75,047
Number of breakout papers detected	24,277	9,020	466	800	5,873
Share of total	1.0%	0.4%	0.1%	0.2%	7.8%
Organisational category					
University (U)	56.5%	69.8%	69.1%	83.6%	56.5%
Research institute (R)	4.4%	6.5%	9.2%	1.4%	3.3%
Company (C)	2.2%	1.6%	4.1%	0.3%	2.0%
Hospital (H)	1.5%	0.0%	0.0%	0.3%	0.4%
U + R	13.8%	15.7%	9.9%	7.9%	21.1%
U + H	8.0%	0.4%	0.2%	3.3%	4.1%
U + C	6.9%	5.7%	7.5%	2.5%	9.2%
U + H + R	3.6%	0.1%	0.0%	0.4%	2.2%
U + H + C + R	1.6%	0.0%	0.0%	0.1%	0.8%
U + H + C	1.5%	0.1%	0.0%	0.4%	0.5%

* Excludes papers that could not be assigned to organisational subcategories

** Journals assigned to this category, by Thomson Reuters, include *Nature*, *Science* and *PNAS*.

Across all fields university staff co-authored between 87% (Engineering Sciences) and 98% (Social and Behavioural Sciences) of the breakout papers. In general the combinations U+R and U+C produce a significant share of breakout papers. For ‘Medical and Life Sciences’, not surprisingly, in 18% of the breakouts there is a hospital affiliation for at least one of the authors. The general journals in the category ‘multidisciplinary journals’ have by far the highest numbers of breakout papers, and the share of breakout papers for these journals is much higher than the overall average share of 0.8% we find. Table 2 contains descriptive statistics for *Nature*, *Science*, and *PNAS*. Based on the share of breakout papers all three

³ Netherlands Observatory of Science and Technology (NOWT)

journals are ranked in the Top-50 of journals with the highest share of breakout papers. Table 3 presents the relative performance in producing breakout papers for different organisational categories. This relative performance is defined as the share of breakout papers for a certain combination of ‘organisational category’ and ‘NOWT-category’ compared to the corresponding share in the total set of papers. In this table a ‘+’ indicates that breakouts are overrepresented, and a ‘-’ that they are underrepresented. Papers in the ‘Engineering Sciences’ from research-institute-affiliated authors, and papers from company-affiliated authors published in a multidisciplinary journal have a higher probability to be a breakout paper. In general papers co-authored by authors of different types of organisations, produce more breakouts than expected. ‘Engineering Sciences’ is an exception to this; in this case only papers with authors from the combination ‘U + R’ have this property.

Table 2 Descriptive statistics for the multidisciplinary journals Nature, Science and PNAS (articles + letters, 2007-2011)

Multidisciplinary journal	Number of papers	Number of breakout papers	Breakout papers as share of total number of papers
<i>Nature</i>	13,041	1,966	15.1%
<i>PNAS</i>	20,173	1,758	8.7%
<i>Science</i>	12,744	1,585	12.4%

Table 3 Distribution of breakouts: over- and underrepresentation relative to all publications in the WoS database (articles + letters, 2007-2011)

	Medical and Sciences	Natural Life Sciences	Engineering Sciences	Social Behavioural Sciences	and Multidisciplinary Journals
Organisational category					
University (U)	-	-	-	-	-
Research Institute (R)	-	-	+	-	-
Company (C)	-	-	-	-	+
Hospital (H)	-	-	-	-	-
U + R	+	+	+	+	+
U + H	+	+	-	+	+
U + C	+	+	-	+	+
U + H + R	+	+	-	+	+
U + H + C + R	+	+	-	+	+
U + H + C	+	+	-	+	+

Size of the author team in relation to the breakout character of a paper

For the selection of papers used in this study we further analysed the data to find out if the average size of the author teams for breakout papers differs from that for papers in general. Figure 1 shows the distribution of the number of papers vs. the size of the author team. This figure shows that the distribution is skewed. For ‘all’ papers the distribution tops at 3 authors, and for the breakout papers at 4 authors. The differences between the distributions are also

illustrated in Table 4 where we present percentile borders for the two distributions. Zooming in on the data reveals that the number of authors that contribute to a paper depends on the organisational categories and on the organisational collaboration. In Table 5 we present the weighted average size of the author team of a paper.

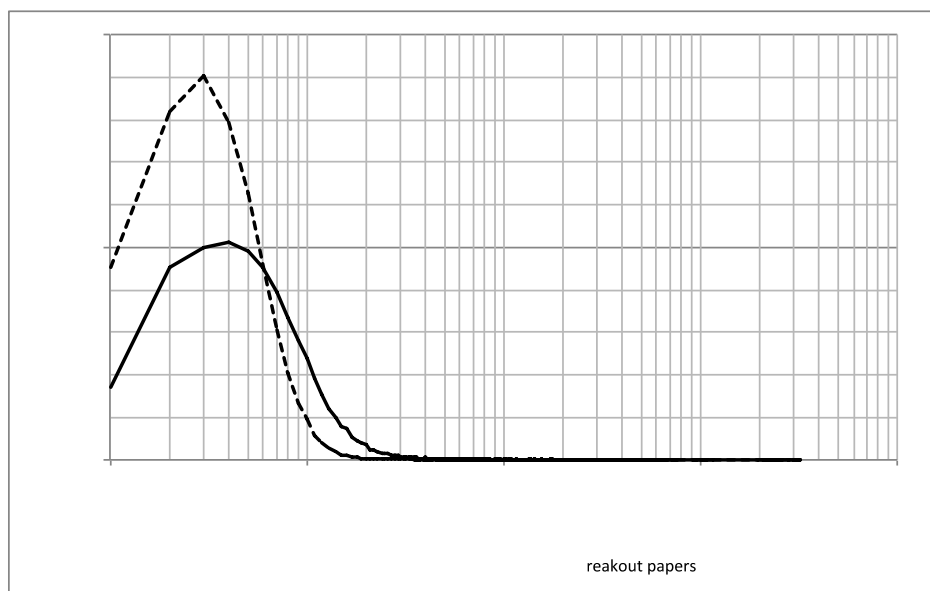


Figure 1 Distribution of the number of authors of a paper for ‘breakout’ and for ‘all’ papers from 2007-2011

Table 4 Percentile borders for the distributions of the numbers of authors for both the ‘breakout’ papers and for ‘all’ papers from 2007-2011

Percentile borders	Breakthrough papers <i>Number of authors</i>	All papers
25%	3.2	2.0
50%	4.8	3.6
75%	9.8	5.3

Table 5 Weighted average number of author per paper of WoS types article and letter from 2007-2011

<i>Breakout papers</i>		<i>All papers</i>	
<i>Organisational category</i>	<i>Average number of authors per paper</i>	<i>Organisational category</i>	<i>Average number of authors per paper</i>
University (U)	6.3	University (U)	4.3
Research Institute (R)	5.8	Research Institute (R)	4.2
Company (C)	8.6	Company (C)	4.9
Hospital (H)	6.2	Hospital (H)	4.4
U+R	16.9	U+R	6.6
U+H	11.1	U+H	6.8
U+C	16.8	U+C	6.8
U+H+R	24.0	U+H+R	12.1
U+C+H+R	43.2	U+C+H+R	19.7
U+H+C	17.4	U+H+C	9.5

On average more authors contribute to breakout papers than to papers in general. On the basis of these results we conclude that particularly breakout papers with one or more authors affiliated to a company have on average a larger author team than papers with only authors affiliated to universities, research institutes or hospitals. The data also shows that collaboration between organisations of different categories leads to substantially more authors. The average size of the author team depends on the organisational collaboration, and is significantly smaller for ‘mono organisational category’ papers. For breakout paper this difference is larger. The author teams for breakout papers are between 1.4 and 2.6 times larger than for all papers.

CONCLUDING REMARKS

Our algorithms for early stage identification of breakout papers use standard the moment of publication as point of reference. We already altered the algorithms so they can analyse a paper’s citation profile to search for breakout characteristics at any moment in time, and not just from the moment of publication. These extended versions of the algorithms reveal that 92% of the publications for which the algorithms discover a breakout characteristic in the first 10 years after publication this characteristic is already visible in the first year after publication.

We find that organisational collaboration influences both the number of breakthrough papers and the average size of the author team of a paper. The results in this study further show that the impact of the organisational collaboration on the production of breakout papers varies across science fields. In multidisciplinary journals the share of papers identified as a breakout paper is highest.

Presented are the preliminary results of research in progress. In the future we will present on more results of this research and extend our typology.

REFERENCES

- Bettencourt, L. M. A., Kaiser, D. I., and Kaur, J. (2009). Scientific discovery and topological transitions in collaboration networks. *Journal of Informetrics*, 3(3):210–221. Science of Science: Conceptualizations and Models of Science.
- Cyranoski, D. (2004). Korea’s stem-cell stars dogged by suspicion of ethical breach. *Nature*, 429(6987):3–3.
- Fleischmann, M. and Pons, S. (1989). Electrochemically induced nuclear- fusion of deuterium. *Journal of electroanalytical chemistry*, 261(2A):301–308.
- Martin, B. R. (1995). Foresight in science and technology. *Technology Analysis & Strategic Management*, 7(2):139–168.
- OECD (1992). Technology and the economy: the key relationships, volume 42 of TEP – The Technology / Economy Programme. Organisation for Economic Co-operation and Development.
- Ponomarev, I. V., Williams, D. E., Hackett, C. J., Schnell, J. D., and Haak, L. L. (2014). Predicting highly cited papers: A method for early detection of candidate breakthroughs. *Technological Forecasting and Social Change*, 81(January 2014):49–55.
- Redner, S. (2005). Citation statistics from 110 years of physical review. *Physics Today*, 58(6):49–54.

- Schneider, J. W. and Costas, R. (2015). Identifying potential breakthrough publications using refined citation analyses: Three related explorative approaches. pages 1–33. Available in arXiv under id 1512.01388v1.
- Smits, R. (2002). Innovation studies in the 21st century: Questions from a user's perspective. *Technological Forecasting and Social Change*, 69(9):861–883.
- Winnink, J. J., Tijssen, R. J. W., and van Raan, A. F. J. (2016). Can early-detection algorithms of breakout papers uncover scientific breakthroughs? (submitted for publication).



From university research to innovation Detecting knowledge transfer via text mining

Sabrina Woltmann^{*}, Line H. Clemmensen^{**} and Lars Alkærsg^{*}

^{*}swol@dtu.dk; lalk@dtu.dk

Management Engineering, Technical University of Denmark, Centrifugevej 372, 2800 Kgs. Lyngby, (Denmark)

^{**}lkhc@dtu.dk

Applied Mathematics and Computer Science, Technical University of Denmark, 2800 Kgs. Lyngby, (Denmark)

ABSTRACT

Knowledge transfer by universities is a top priority in innovation policy and a primary purpose for public research funding, due to being an important driver of technical change and innovation. Current empirical research on the impact of university research relies mainly on formal databases and indicators such as patents, collaborative publications and license agreements, to assess the contribution to the socioeconomic surrounding of universities. In this study, we present an extension of the current empirical framework by applying new computational methods, namely text mining and pattern recognition. Text samples for this purpose can include files containing social media contents, company websites and annual reports. The empirical focus in the present study is on the technical sciences and in particular on the case of the Technical University of Denmark (DTU). We generated two independent text collections (corpora) to identify correlations of university publications and company webpages. One corpus representing the company sites, serving as sample of the private economy and a second corpus, providing the reference to the university research, containing relevant publications. We associated the former with the latter to obtain insights into possible text and semantic relatedness. The text mining methods are extrapolating the correlations, semantic patterns and content comparison of the two corpora to define the document relatedness. We expect the development of a novel tool using contemporary techniques for the measurement of public research impact. The approach aims to be applicable across universities and thus enable a more holistic comparable assessment. This rely less on formal databases, which is certainly beneficial in terms of the data reliability. We seek to provide a supplementary perspective for the detection of the dissemination of university research and hereby enable policy makers to gain additional insights of (informal) contributions of knowledge dissemination by universities.

INTRODUCTION

Universities are facing increasing demands for active dissemination of their research results and their contribution to knowledge development in their socioeconomic environment (Jongbloed, Enders & Salerno 2008); commonly referred to as the third mission. Since knowledge is a crucial aspect for innovation processes, its transfer has become a new policy priority and is often directly targeted by public funding (Ramos-Vielba, Fernández-Esquinas & Espinosa-de-los-Monteros 2009, Huggins & Johnston 2009). This study covers the extent

of knowledge transfer of university research within technical sciences, as these are key drivers for innovation.

Current empirical research focuses primarily on the analysis of formal interactions between universities and their company partners (Broström 2010), relying on indicators such as patents, collaborative publications, contracts and license agreements (Drucker & Goldstein 2007). These well-developed empirical approaches somewhat capture the success of knowledge dissemination and commercialization of university driven innovations. However, these studies bear some deficiencies, as they often fail to include indirect impacts by focusing on formal cooperation and knowledge exchange. Additionally, most empirical studies frequently require complex adjustments for each unique case. Moreover, their key indicators often depend on formal databases with varying quality and accessibility and they require long-term assessments, which delays the outcomes and limit comparability (Vincett 2010).

In this study, we use modern computational methods to expand the empirical framework by introducing specific data mining approaches and testing these on the Technical University of Denmark (DTU). To complement the current scope, we focus in particular on the application of text mining and pattern recognition tools. These tools capture occurrences where knowledge is used without a statement about its origin. Our data sources include the online presence of companies in regional proximity to the university, including social media sites, company websites and annual reports.

The study's intent is to counteract certain empirical challenges, by detecting knowledge transfer without focusing on formal cooperation channels and develop additional indicators also capturing informal contributions. Compared to traditional assessments, the main advantages are that the measure is instantaneous, resulting in reduced time delay, and that it relies less on formal databases.

METHODOLOGY

We seek to generate a complementary perspective by applying novel computational methods and embedding them in the current impact assessment framework of public research, therefore seizing the widely agreed potential of those applications by adapting them to our specific purpose. We capture, identify and verify the existence and extent of knowledge transfer to the economic surrounding of the DTU. To identify research outcomes, which can be attributed to the university, this study uses text-mining methods.

To implement these measures we follow systematic and distinct actions, including the sample generation, data collection, pre-processing and the application of statistical correlation measures.

Sample generation

Assuming that the private economy is an essential beneficiary of knowledge exchange, we included relevant private companies. Defined as companies

- with direct relations to DTU defined by hyperlinks on the DTU website (first-degree partners) ;
- with indirect relations to DTU defined by hyperlinks on partner websites (second-degree partners);

- with regional facilities near by the DTU (in the national context of Denmark, indicated via a Danish VTA registration).

Data collection

This study uses company websites, which consist of unstructured text-data, to identify company knowledge, products and expertise. Thus, we gathered these texts, which are available in form of online publications released by the companies themselves, and extracted them as HTML files. Associated social media entries will be included at a later stage of the project, as social media content requires specific treatment due to their special linguistic composition. The collected HTML files are pre-processed and transformed into unstructured raw text, maintaining only content and semantically relevant information. We implemented language identification parameters to extract exclusively English texts (Palmer 2010).

Text mining

To analyze the data, we apply methods from the field of natural language processing (NLP), as it provides tools for simple and advanced text analytical procedures. Text mining requires text corpora containing the relevant text fragments in form of tokens. In our case, we developed one text corpora derived from the raw text files of the company websites and a second 'reference' corpus containing an extensive sample of research publications. The university online publication database ORBIT provided the texts for the reference corpus, as this database comprises almost all publications including patents, projects, etc. made by DTU employees¹.

Pattern recognition algorithms and machine learning methods provide in-depth comparisons between the reference and the company corpus (Bird, Klein & Loper 2009). To extrapolate the important patterns, including correlations, semantic compositions and outlier comparison, this study uses various available text mining methods. These include term-based methods, phrase-based methods, etc., which provide a variety of statistical tools to analyze the texts and to achieve our objectives. The analysis includes statistical measures that identify document relatedness, correlations or different types of regression parameters. Hereby, we quantify the extent of correlations between documents of the two corpora and the corpora themselves.

To detect the similarities between texts from the two corpora we use specially adapted machine learning algorithms, which extract key features from the reference corpus and compare them with the company corpus. We aim to include semantically correlated and content related approaches, to ensure the methods capture not only obvious semantic, but also content correlations. Accordingly, this approach allows us to detect shared contents among documents and enables the tracing of knowledge, which provides evidence-based insights in the 'relatedness' between the corpora.

We use statistical models, which include, but are not limited to, methods for dimensionality reduction like latent semantic analysis (LSA) (Landauer, Foltz & Laham 1998) and, for uncovering the underlying structures of the documents, probabilistic topic models for instance latent dirichlet allocation (LDA) and correlated topic models (CTM) (Blei, Ng & Jordan 2003). However, as NLP is a comparatively young field its methods undergo continuous development, therefore specific adjustments to its models are inevitable.

¹ <http://orbit.dtu.dk/en/about.html>

Evaluation of the method

Given the identification of the extent of knowledge transfer by tracing linguistic and semantic content, we seek to extrapolate the research areas, which spread most knowledge and the companies, which make most use of university research within their proximity. To evaluate the relevance of our findings and to conclude whether our findings truly increase the understanding and measurement of (indirect) impacts we will compare our results to those of conventional measures.

POLICY RELEVANCE AND POTENTIAL

The study provides a supplementary perspective for the detection of research dissemination and impact of university innovations. Our intention is to contribute to the understanding of university performance by enhancing the detection of impacts of publicly funded research. Current computational methods provide novel possibilities for measurements allowing additional benchmarking as foundation for decision-making processes.

The goal is to provide policy makers with additional insights on the applications of university knowledge, allowing them to evaluate the benefits of government funding of research in a more holistic manner by including so far undetected, but essential impacts. This study can shed new light on the contributions universities make to economy and society.

Advantages of this novel approach are firstly, the availability of data, contrary to conventional assessments, which rely highly on university databases, which vary in quality and accessibility. Secondly, the potential to apply these measures in different regional, societal and economic contexts. Thirdly, the instantaneous nature of the measurement could capture the outcomes and the status quo almost in real-time.

After an in-depth evaluation of our approach against existing measures, we will be able to verify the extent of additional information that can be drawn from this new approach. Ideally, it will provide a greater overview about (informal) knowledge exchange from universities to companies, providing a more detailed picture for future oriented decision-making.

REFERENCES

- Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python*. Sebastopol: O'Reilly Media, Inc.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan), 993-1022.
- Broström, A. (2012). Firms' rationales for interaction with research universities and the principles for public co-funding. *The Journal of Technology Transfer*, 37(3), 313-329.
- Drucker, J., & Goldstein, H. (2007). Assessing the regional economic development impacts of universities: a review of current approaches. *International regional science review*, 30(1), 20-46.

- Gaikwad, S. V., Chaugule, A., & Patil, P. (2014). Text mining methods and techniques. *International Journal of Computer Applications*, 85(17).
- Huggins, R., & Johnston, A. (2009). The economic and innovation contribution of universities: a regional perspective. *Environment and Planning C: Government and Policy*, 27(6), 1088-1106.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse processes*, 25(2-3), 259-284.
- Palmer, D. D. (2010). *Handbook of natural language processing (second edition)*. Boca Raton: CRC Press.
- Jongbloed, B., Enders, J., & Salerno, C. (2008). Higher education and its communities: Interconnections, interdependencies and a research agenda. *Higher education*, 56(3), 303-324.
- Ramos-Vielba, I., Fernández-Esquinas, M., & Espinosa-de-los-Monteros, E. (2009). Measuring university–industry collaboration in a regional innovation system. *Scientometrics*, 84(3), 649-667.
- Vincett, P. S. (2010). The economic impacts of academic spin-off companies, and their implications for public policy. *Research Policy*, 39(6), 736-747.



Predicting panel scores by linguistic analysis*

Peter van den Besselaar, Luka Stout, Xiaoli Gou

p.a.a.vanden.besselaar@vu.nl

Network Institute and Department of Organization Sciences
Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

ABSTRACT

In this paper we explore the use of text analysis for deriving quality indicators of project proposals. We do full text analysis of 3030 review reports. After term extraction, we aggregate the term occurrences to linguistic categories. Using these linguistic categories as independent variables, we study how well these predict the grading by the review panels. Together, the different linguistic categories explain about 50% of the variance in the grading of the applications. The relative importance of the different linguistic categories inform us about the way the panels work. This can be used to develop altmetrics for the quality of the peer and panel review processes.

INTRODUCTION

In panel review, panel members are asked to apply more or less explicitly specified criteria. Whether they do, and how they score and weight these criteria remains invisible and under-investigated. What is in principle accessible for investigation are the review reports the applicants receive. The question that comes up is if we can use the review reports to (i) identify the evaluation dimensions that play a role, and (ii) identify the criteria deployed and their weight in the final decisions.

Full text analysis has become increasingly powerful in analyzing evaluative texts, such as reports of annual performance interviews (Semin & Fiedler 1991; Bienat et al 2012; Kaatz 2012a). The main focus in that research is on the presence of bias in decision-making, such as gender bias (Bienat et al 2012; Kaatz 2014b). Indeed, quite some empirical support is available for the gender stereotyping practices in hiring and promotion in general, but also in hiring, promoting, and grant decisions within science (Millar et al 2014). This study was also motivated by the question of gender bias in grant allocation. However, we feel that it is worthwhile to investigate the quality of peer and panel review, and of grant decision making *in general* using text analysis.

Data and methods

In practice this is a difficult task, as researchers generally do not have access to these review reports – especially not to the review reports of the rejected applications.¹ In the context of a project on gender bias, we in fact do have the review reports of about 3030 applicants, of which 352 successful and 2674 rejected. We also have the final scores for the applicants: one

* The support of the ERC (grant 610706, the GendERC project) and of the Network Institute, Vrije Universiteit Amsterdam (Academy Assistants project) is acknowledged. For useful comments, we thank Ulf Sandström (KTH, Stockholm, Sweden) and Charlie Mom (Amsterdam, the Netherlands), Helene Schiffbacker, Florian Holzinger (both Joanneum Research, Vienna, Austria) and Claartje Vinckenburg (Vrije Universiteit, Amsterdam, The Netherlands).

for the PI and one for the project in step 1 leading to a decision whether the applicants can go to step 2 (about 25%), or is out (about 75%). We have the scores obtained in step 2, determining whether one gets the grant or not (both groups about 50%).

The 3030 review reports consist of between 4 and 13 short individual reviews. The analysis presented here combines them into a single document. Pre-processing was done to remove the irrelevant text: instruction sentences, section headings, etc. We automatically processed some samples, and checked manually whether the conversion from PDF into TXT worked correctly.

We used LIWC (<http://liwc.wpengine.com/>), a tool for linguistic analysis of texts, based on a variety of predefined linguistic categories. Each category consists of a series of words representing that category, which were validated in other studies (Abele and Wojciszke 2014). The LIWC program counts for each of the categories how many times a word belonging to that category is present in a review report. As the reports are of different length, normalization is needed: the number is translated into a percentage. We first selected those categories used in previous studies on grant decision and panel deliberation (Kaatz et al. 2014a):

- *Ability* words, such as gift*, intell*, skill*;
- *Achievement* words such as creati*, excel*, compet*; (and negative achievement words as separate category).
- *Agentic* words such as outspoken, solid, risk;
- *Negative evaluation* words such as naïve, defect*, lack*;
- *Positive evaluation* words such as intriguing, compelling, commit*;
- *Research* words such as laboratory, result*, fund*;
- *Standout adjectives* such as world class, outstanding, exceptional*.

CorTexT (<http://cortext.risis.eu/login>) was used for term extraction from the review reports. The list of 10.000 most frequently used (stemmed) terms was inspected in order to find additional evaluative terms from the reviews. We added these terms to the LIWC dictionary: Ability (10% extended); Achievement (20% extended); Agentic (0% extended²); Negative evaluation (25% extended); Positive evaluation (30% extended); Research (9% extended); Standout adjectives (3% extended). Also based on the term extraction, we decided that some other categories should be used (without extending):

- *Negating*³ words such as hasn't, don't, can't;
- *Negative emotions* words such as abuse*, bitter*, bad*;
- *Positive emotions* words such as agreeabl*, benefit, helpful;
- *Exclusion* words such as but, either, except, just, not;
- *Insight* words such as define, reflect, idea*;
- *Certainty* words such as fundamental, commitment, truly.

Why were these selected? For *negation* words the argument goes as follows: the excellent applicants are the norm in science, and the other are measured against those excellent: „not excellent“. One would expect much more negation words in evaluation reports of non-successful applicants than in those of successful applicants. *Exclusion* words might be used biased because of the same argument. Positive and negative *emotions* are relevant to include, as one would want to see how strong *sentiments* play a role in panel deliberation.

We also added a few other categories, not from LIWC, but related to the specific evaluation practice under study. „Research“ words (one of the categories mentioned above) relate to the track record of the PI and to the quality of the research proposal. As these are two different

evaluation categories in the case under study, we have split „research“ into „track record“ and „proposal“. We also extended the „standout“ category with a series of typical ERC superlatives. In the analysis below, „superlatives“ replaces „standout adjectives“, and „track record“ and „proposal“ replace the „research“ category.

- *‘Superlatives’* such as groundbreaking, grand challenge, forefront, great potential, high risk high gain;
- *Track record* such as high impact journal, coauthor, H-index, editor, advisor
- *Proposal* words such as multidisciplinary, timeline, laboratory.
-

Using the category-ratios obtained by LIWC, we first compare the language used in review reports between those accepted and those rejected in step 1, and between the granted and non-granted applicants in step 2. This will inform us which word categories are related to success and to the opposite. Then we deploy linear regression to predict the four scores from the linguistic categories. From this we may learn what quality dimensions are relevant and in what order of importance.

Success versus no success

In table 1, we show the ratio of *average occurrence* of the word categories in the different groups: those that were accepted versus those that were rejected in step 1; and the granted versus non-granted applicants in step 2.

Review reports of applicants that are accepted in step 1 and therefore proceed to step two show *significantly more* certainty words, superlatives, agentic words, ability words, achievement words, positive evaluation words, and *significantly less* negating words, negative evaluation words, exclusion words, track record words, negative achievement words, research words, proposal words, insight words (Table 1, left side). Comparing in step 2 the granted versus the rejected applications, we find the same pattern, with only one difference: positive and negative emotion words make a difference in step 2 but not in step 1.

Table 1: Relative frequency of word category use

Accept/Reject Step 1*	Ratio***	Sig.	Granted/Rejected Step 2**	Ratio***	Sig.
Certainty	1.20	0.000	Certainty	1.18	0.000
Agentic	1.08	0.000	Superlatives	1.10	0.000
Ability	1.08	0.002	Agentic	1.09	0.000
Superlatives	1.05	0.002	Ability	1.09	0.017
Achievement	1.04	0.000	Achievement	1.06	0.000
Evaluation positive	1.04	0.000	Evaluation positive	1.06	0.000
Emotions positive	1.00	0.872	Emotions positive	1.03	0.073
Emotion negative	0.98	0.440	Emotion negative	0.94	0.060
Insight	0.96	0.000	Insight	0.96	0.006
Proposal	0.92	0.000	Proposal	0.94	0.000
Research	0.89	0.000	Research	0.91	0.000
Achievement negative	0.87	0.010	Achievement negative	0.79	0.003
Track record	0.75	0.000	Track record	0.78	0.000
Exclusion	0.73	0.000	Exclusion	0.70	0.000
Evaluation negative	0.68	0.000	Evaluation negative	0.64	0.000
Negate	0.57	0.000	Negate	0.53	0.000

ERC Starting grants 2014, 3030 applicants

* *N = 785 resp. 2241*

** *N = 352 resp. 2674 (four reports missing)*

*** *Ratio = ratio of the mean word frequency of the two groups, and the significance level comes from an Anova test. Ratios > 1 mean that the linguistic category occurs more in review reports about those that proceeded to step 2 (left) and granted applicants (right); ratios < 1 mean that the linguistic category occurs more often in review reports about applicants that were rejected in step 1 (left) and non-granted applicants (right).*

Intuitively this seems a reasonable result. On the positive side we see the positive categories (positive achievement words; positive evaluation words, positive emotions, standout adjectives and ERC terms), and the categories pointing at certainty, strong agency and ability. On the negative side we found the direct negative categories (negating and negative evaluation words, exclusion words, negative achievements, and negative emotions). Furthermore, we find that if track record, insight and the research proposal are discussed, it is generally at the negative side. And, emotions do not seem to play a role in the first step, which suggest a more rational decision-making process, but they do in step 2.

Does this tell us what criteria are used? To some extent: Achievements and performance of the applicants (achievement words, track record words), personal characteristics (agentic words, ability words), excellence of the applicants (superlatives), quality of the proposal (proposal words) and evaluations by the panelists (evaluation words, negation words, emotion words). Interestingly, these are *dimensions* relevant in decision-making, and not operational selection criteria. In fact, in the case under study no specific criteria are formulated for the reviewers – they only should look at „excellence“. This may imply that panelists have different understanding about how to evaluate. We indeed do not find terms representing specific criteria consistently used through the reviews.

What evaluation dimensions are important?

In this section, we predict the two scores in step 1 (step1-PI score and step1-Project score) and the scores in step 2 (step2-PI score and step 2-project score) by the frequencies of the linguistic categories as the independent variables. As this is an explorative study, we deploy linear regression with a „stepwise“ inclusion of the independent variables. This leads to a model in which the non-significant variables are left out of the model (Table 3). As words may be used not independently, we tested for multicollinearity. Tolerance and VIF values suggest that this does not occur.

We start with the prediction of the PI-score in step 1. Table 2 shows the linguistic categories that remain in the (final) model. We report here the standardized regression coefficients, enabling to assess the effect-size of the variables when controlling for the effect of the others.

Table 2: PI-score in step 1 by linguistic categories

Model 12	Unstandardized Coefficients		Standardized Beta	t	Sig.	Collinearity Statistics	
	B	Std. Error				Tolerance	VIF
(Constant)	3.211	0.069		46.721	0		
Negate	-0.446	0.027	-0.463	-16.625	0	0.225	4.446
Negative Evaluation	-0.172	0.019	-0.218	-8.973	0	0.295	3.388
Certainty	0.195	0.022	0.121	8.776	0	0.913	1.095
Proposal	-0.04	0.007	-0.083	-5.998	0	0.911	1.098
Positive evaluation	0.044	0.011	0.080	3.939	0	0.419	2.387
Superlatives	0.037	0.013	0.056	2.867	0.004	0.463	2.161
Track record	-0.037	0.010	-0.057	-3.801	0	0.783	1.277
Exclusion	0.041	0.019	0.052	2.136	0.033	0.294	3.399
Negative achievements	-0.145	0.055	-0.037	-2.627	0.009	0.881	1.135
Agentic	0.036	0.015	0.034	2.467	0.014	0.903	1.107
Positive emotions	0.032	0.012	0.040	2.671	0.008	0.760	1.315
Insight	0.021	0.011	0.028	2.011	0.044	0.927	1.078

R-square = 0.485

As table 2 shows, the variables with the strongest effects are those that are negatively associated with the score: negation words, and negative evaluation words. Also words referring to the research proposal and to the track record have a negative correlation with the score, suggesting that when the panel talks about the proposed project and the track record, this is more often in a negative than in a positive way. Panel discussions seem more about talking proposals down. The other categories have significant positive relations with the score, but the effect of most is rather small. Interesting is that achievements and negative emotions have no effect, and the positive emotions only a very small one. Table 3 lists the variables not in the analysis. The final model has an R-square of .485, which means that the variance is almost half explained by the linguistic variables.

Table 3: Variables not used in the analysis

Model 12	Beta In	t	Sig.	Partial Correlation
Negative emotions	-.004	-0.313	0.755	-0.006
Achievement	-.021	-1.304	0.192	-0.024
Ability	-.013	-0.911	0.363	-0.017

We now turn to the score for the Project (also in step 1 of the procedure). The results are in table 4, and these are similar to the results for the PI in step 1.⁴ A main difference is that related to the project, also *negative emotions* play some role. The explained variance is quite high (0.521), even higher than in the previous analysis.

Table 4: PROJECT-score in step 1 by linguistic categories

Model 11	Unstandardized Coefficients		Standardized Beta	t	Sig.	Collinearity Statistics	
	B	Std. Error				Tolerance	VIF
(Constant)	3.219	0.054		59.607	0		
Negate	-0.469	0.024	-0.522	-19.834	0	0.231	4.327
Negative Evaluation	-0.173	0.016	-0.234	-10.584	0	0.327	3.059
Certainty	0.154	0.020	0.102	7.730	0	0.915	1.093
Exclusion	0.071	0.017	0.097	4.164	0	0.295	3.394
Proposal	-0.039	0.006	-0.088	-6.645	0	0.914	1.094
Superlatives	0.053	0.012	0.085	4.602	0	0.466	2.148
Positive evaluation	0.043	0.010	0.083	4.268	0	0.421	2.378
Track record	-0.034	0.009	-0.055	-3.885	0	0.790	1.267
Positive emotions	0.033	0.011	0.044	3.066	0.002	0.761	1.314
Negative emotions	-0.034	0.015	-0.030	-2.333	0.02	0.960	1.042

R-square = 0.521

Table 5: Variables not used in the analysis

Model 11	Beta In	T	Sig.	Partial Correlation
Insights	-.005	-0.394	0.694	-0.007
Achievement	.004	0.310	0.757	0.006
Ability	.005	0.411	0.681	0.007
Agentic	.015	1.136	0.256	0.021
Negative achievements	-.002	-0.130	0.897	-0.002

CONCLUSIONS

This methodological experiment suggests that word use in review reports can inform us about the relevance of specific evaluation dimensions in grant decision-making. A few interesting observations were made. The strongest effect comes from the negative linguistic categories: negation words and negative evaluations. This suggests that panels are concentrate on finding those proposals that in their view should be deleted from the procedure, and not on the best one to get funded. Where panels discuss more intensively the track record of the PI and the quality of the proposal, it is more often in a negative than in a positive way. The explained variance of the models is fairly high. Overall, the results suggest that further exploration of the approach is promising.

FURTHER WORK

In a next step we will analyze step 2, where the success rate is much higher than in step 1. If everything has worked well, the applications in step 2 all of high quality – which means a different way of argumentation in the selection process. So one would expect that other linguistic dimensions become dominant.

In this paper, we use the review reports of an application as *one text*. In the next stage will use the different reviews as separate documents. This will enable us to include the variety of language use between the reviewers of the same application. Another extension is to do the

analysis separately for the different parts of the review report: the assessment of the (groundbreaking) contribution of the project proposal, of the (methodological) feasibility of the project, and of the background (performance) of the applicants.

Furthermore, in this paper we use mainly monograms and bigrams. We will test the approach also using n-grams. Another aspect that might be studied is the context in which the negation words are used.

As grading and decision-making takes place within panels, we may add this level in multi-level analysis of the data.

It also would be highly relevant to add other variables to the analysis, especially those reflecting *peripheral spaces* of the science system: if we enter gender, would we find different assessment dimensions being dominant for male and female applicants, e.g., those reflecting gender stereotyping (Miller et al 2014; Kaatz et al 2014b)? And if we would distinguish between applicants from core and from peripheral countries, what would that show? As the data are available, this can relatively easily be done (but not within the space limits of this paper).

Finally, one may apply lexical analysis also to CVs and to the project descriptions. Does the language used (sentiments and other linguistic dimensions) relate to success?

REFERENCES

- Biernat M, Tocci MJ, Williams JC, The language of performance evaluations: gender-based shifts in content and consistency of judgment. *Social Psychological and Personality Science* 3 (2012) 186-192
- Buttlere B, *Identifying high impact scientific work using natural language processing and psychology*. Paper workshop “Quantifying and Analysing Scholarly Communication on the Web (ASCW‘15)”
- Kaatz A, Magua W, Zimmerman DR, Carnes M, A quantitative linguistic analysis of National Institutes of Health R01 application critiques from investigators at one institution. *Academic Medicine* (2014a)
- Kaatz A, Gutierrez B, Carnes M, Threats to objectivity in peer review, the case of gender. *Trends in Pharmacological Sciences* 35 (2014b) 371-373
- Miller DI, Eagly AH, Linn MC, Women’s representation in science predicts national gender-science stereotypes: evidence from 66 nations. *Journal of Educational Psychology* (2014)
- Semin GR, Fiedler K, The linguistic category model, its bases, applications and range, *European Review of Social Psychology* 2 (1991)

¹ Even more difficult than getting review reports is getting access to observe panel discussion and decision-making. In the GendERC project we tried to do this, but it proved to be impossible.

² To be tested: to what extent should we split the categories into positive and negative? E.g., masculine is an *agentic* word, but energetic is an *agentic* word too. We have done such a test for positive versus negative *achievement words*. We also test whether splitting *research* into *proposal* and *track record* is relevant and informative. The latter is based on the distinction between evaluating the *Principle Investigator* and the *Project*. In both cases (achievement and research), splitting the category does not make much difference.

³ Kaatz et al. (2014b) explain why they did not include *negation words*: it did not correlated with one or some of the other categories, so therefore it was not needed to include it as negative qualification of some of the other categories. However, this is different in our much larger set, as the category of negating words correlates very strong (>.9) with negative evaluations.

⁴ This is in line with the finding that the scores for project and for track record correlate very strong.



Networks dynamics in the case of emerging technologies¹

Daniele Rotolo*

*d.rotolo@sussex.ac.uk

SPRU — Science Policy Research Unit, University of Sussex, Brighton, BN1 9SL, United Kingdom

ABSTRACT

This research in progress aims at increasing our understanding of how collaborative networks form, evolve and are configured in the case of emerging technologies. The architecture of the relationships among the variety of organisational actors involved in the emergence process exerts a significant influence in shaping technological change in certain directions rather than others, especially in the early stage of emergence. As a result, socially optimal or desirable technological trajectories may be ‘opportunistically’ rejected. Our empirical analysis is based on a case-study of an emerging medical technology, namely ‘microneedles’. On the basis of co-authorship data reported in 1,943 publications on the topic from 1990 to 2014, longitudinal collaboration (co-authorship) networks were built at two levels: affiliation and author. We examined the dynamics of co-authorship networks by building on recent methodological advancements in network analysis, i.e. Exponential Random Graph Models (ERGMs). These models enable us to make statistical inferences about on the extent to which a network configuration occurs more than could be expected by chance and to identify which social mechanisms may be shaping the network in certain configurations. The findings of the statistical analyses (currently in progress) combined with the qualitative understanding of the case will increase our understanding of which mechanisms are more likely to drive the network dynamics in the case of emerging technologies. These include evidence of the extent to which the likelihood of forming, maintaining, or terminating ties among actors (authors or affiliations) is affected by actors’ covariates such as types of organisations, diversity/specialisation of the research undertaken, and status. These findings have potential to provide important inputs for policymaking process in the case of emerging technologies.

INTRODUCTION

This *research in progress* aims at increasing our understanding of how collaborative networks form, evolve and are configured in the case of emerging technologies. Emerging technologies are technologies with the potential to exert a considerable socio-economic impact in the domain in which they emerge. They are radically novel, have already moved beyond the conceptual stage, and show relatively fast growth in terms of actors involved in knowledge production processes and outcomes of these processes (e.g. publications, patents, products/services) (Rotolo, Hicks, & Martin, 2015).

Nonetheless, uncertainty and ambiguity exist on the directionality, application, and success of emerging technologies, thus on change to *status quo* these promise (Rotolo et al., 2015). Within this context, networks represent critical conduits for the exchange of knowledge and resources among the different actors (Chesbrough, 2003; Padgett & Powell, 2012) as well as

¹ This paper is very preliminary. Please do not cite it without the consent of the author. This work is supported by the People Programme (Marie Curie Actions) of the European Union’s Seventh Framework Programme (FP7/2007-2013) (award PIOF-GA-2012-331107 - [NET-GENESIS: Network Micro-Dynamics in Emerging Technologies](#)).

loci ‘distributing’ power among actors and enabling actors to build a reputation that extends beyond their peers (signalling effect) (Gulati & Gargiulo, 1999).

As a consequence, the architecture of the relationships among the variety of organisational actors involved in the emergence process exerts a significant influence in shaping technological change in certain directions rather than others, especially in the early stage of emergence (e.g. Collingridge, 1980; Geels, 2002; van Lente & Rip, 1998). For example, a network in which the structure tends to concentrate power over the control of information, knowledge and resources in a limited number of actors (e.g. a few central actors or actors performing as brokers) may create the conditions for the few powerful actors to opportunistically leverage the whole network by controlling the flows of knowledge and resources (Bonacich, 1987; Burt, 1992). As a result, socially optimal or desirable technological trajectories may be ‘opportunistically’ rejected (Stirling, 2009).

Previous studies have extensively investigated the consequences of network variables on actors’ behaviour and performance, stressing the importance of networks to gain social, institutional and governance benefits and private advantages (Burt, 1992; Granovetter, 1983). Yet, the genesis and dynamics of networks is a largely unexplored area of research (Ahuja, Soda, & Zaheer, 2012; Gulati & Gargiulo, 1999). This is especially important in the context of emerging technologies. Given that these are in a state of flux, the architecture of networks (and the distribution of the benefits and advantages among the actors involved) is likely to change over the emergence process. This *research in progress* aims to fill this gap by conducting an in depth case-study analysis that combines bibliometrics and network modelling with interviews.

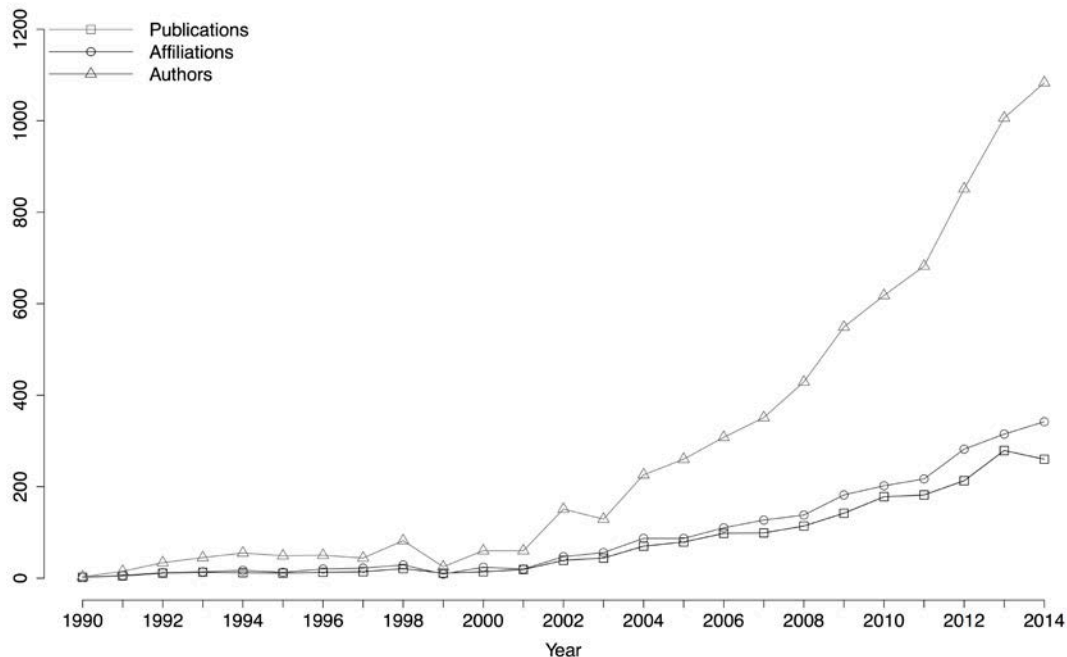
METHODS

The empirical analysis is based on a case ‘microneedles’ a relatively recent emerging technology. The analysis is focused on the understanding of the collaboration network dynamics, which are of a particular importance in the medical context given the key role of networks in the ‘problematisation’ of a technology (Blume, 1992) and the persistent uncertainty in all the stages of the medical innovation process (e.g. Consoli, Mina, Nelson, & Ramlogan, 2016; Petersen, Rotolo, & Leydesdorff, 2016).

Case-study: Microneedles technology

Microneedle are needles the size of which (e.g. diameter, length) is on the micrometer length scale. These are combined in patch-like structures that, when applied on the skin, create painlessly micro-holes through which macromolecular drugs (e.g. vaccines, insulin) can be delivered. This radically novel approach was proposed in the 1970s, but demonstrated in the 1990s thanks to the advancements in microelectronic industry. Considerable expectations on the use of microneedles for vaccination purposes (Koutsonanos et al., 2012) and for drug delivery (Brambilla, Luciani, & Leroux, 2014) exist as well as on the possibility to reduce the production of biohazard waste with their use (microneedles can be designed to dissolve on the skin). Yet, these contrast with persisting uncertainty and ambiguity across a number of dimension including approaches to deliver drugs, materials used to produce microneedles, and safety (e.g. skin irritation).

Figure 1: Number of publications, affiliations, and authors involved with microneedles technology (1990-2014).



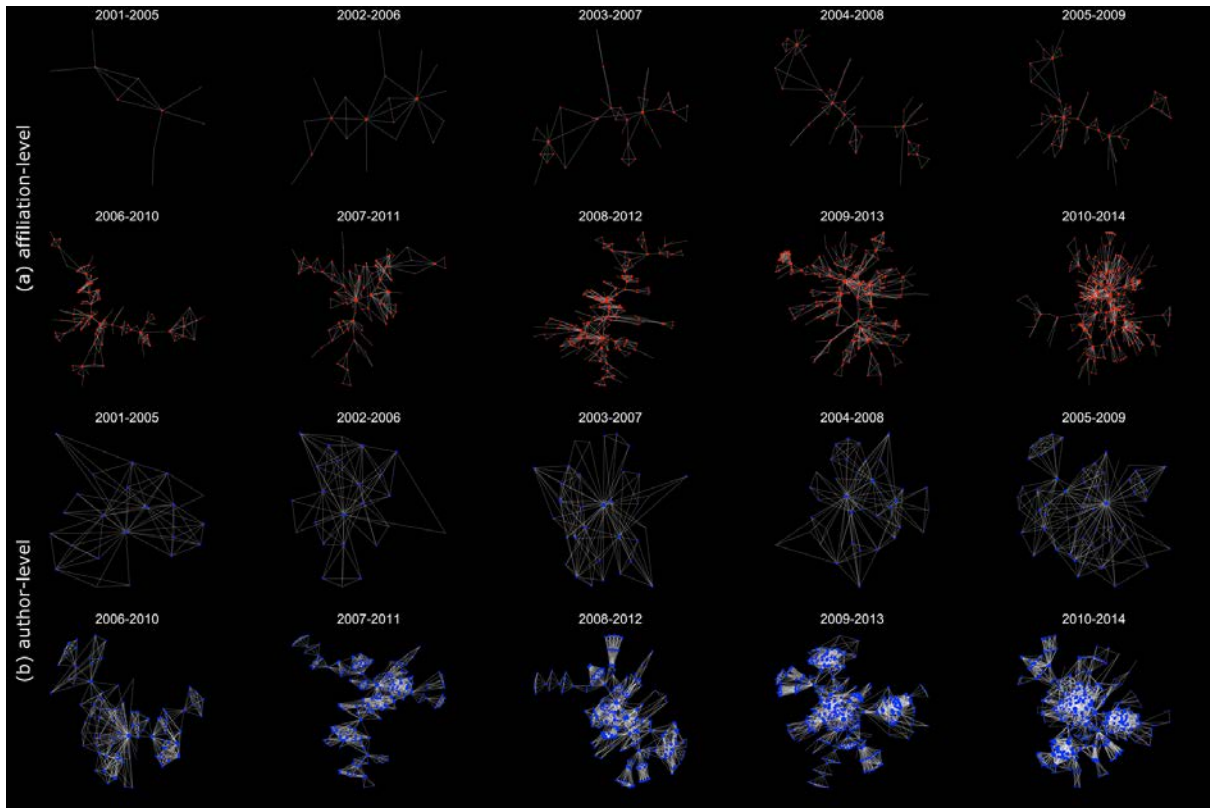
Source: Author's elaboration

Data

Collaboration networks in the context of research on microneedles technology were defined on the basis of co-authorship data reported in publications on the topic. We queried SCOPUS with a list of relevant keywords – defined with the support of experts in the field.² These were searched in publications' titles, abstracts, and lists of keywords reported by authors. The search returned 1,943 publication records from 1990 to 2014. Affiliations and authors' names reported in this sample of publications were then harmonised: 1,240 and 5,164 distinct affiliations and authors were identified, respectively. Figure 1 depicts the number of publications, affiliations, and authors involved with research on microneedles from 1990 to 2014. Co-authorship data were then used to build longitudinal collaboration (co-authorship) networks at two levels: affiliation and author (see Figure 2). Table 1 reports basic network descriptive statistics of the two networks.

² We queried SCOPUS on 10 March 2015 by using the following search string: "TITLE-ABS-KEY(microneedle*) OR TITLE-ABS-KEY(micro-needle*) OR TITLE-ABS-KEY("microprojection patch*") OR TITLE-ABS-KEY("micro-projection patch*") OR TITLE-ABS-KEY("microprojection array") OR TITLE-ABS-KEY("micro-projection array") OR TITLE-ABS-KEY("micromechanical piercing structure*") OR TITLE-ABS-KEY("micro-mechanical piercing structure*") OR TITLE-ABS-KEY("microscopic needle*") OR TITLE-ABS-KEY("micron-scale needle*")".

Figure 2: Longitudinal co-authorship networks (largest component) at affiliation (a) and author (b) levels for the 2001-2014 period.



Source: Author's elaboration

Model

Most conventional statistical models assume independencies among observations. This assumption, however, is violated when using networks data (by definition ties are relational data, thus they are non-independent). For this reason, we examined the dynamics of co-authorship networks by building on recent methodological advancements in network analysis, i.e. Exponential Random Graph Models (ERGMs) (Krivitsky & Handcock, 2014; Robins, Pattison, Kalish, & Lusher, 2007). These models enable us to make statistical inferences about on the extent to which a network configuration occurs more than could be expected by chance and to identify which social mechanisms may be shaping the network in certain configurations. We used the “statnet” package in R (Handcock, Hunter, Butts, Goodreau, & Morris, 2008) to estimate the models.

Table 1. Basic network descriptive statistics.

Affiliation-level					
<i>Time window</i>	<i>1990-1994</i>	<i>1995-1999</i>	<i>2000-2004</i>	<i>2005-2009</i>	<i>2010-2014</i>
<i>Nodes</i>	30	74	123	314	703
<i>Ties</i>	7	54	101	338	1063
<i>Density</i>	0.02	0.02	0.01	0.01	0.00
<i>Average path length</i>	1.00	1.14	1.15	3.94	5.24
<i>Nodes in the largest component</i>	3 (10%)	6 (8%)	6 (5%)	61 (19%)	299 (42%)
<i>Number of components</i>	1	9	15	26	41
Author-level					
<i>Time window</i>	<i>1990-1994</i>	<i>1995-1999</i>	<i>2000-2004</i>	<i>2005-2009</i>	<i>2010-2014</i>
<i>Nodes</i>	93	213	348	1079	2509
<i>Ties</i>	145	451	821	3018	8490
<i>Density</i>	0.03	0.02	0.01	0.01	0.00
<i>Average path length</i>	1.03	1.10	1.27	1.69	3.53
<i>Nodes in the largest component</i>	8 (9%)	14 (7%)	17 (5%)	62 (6%)	420 (17%)
<i>Number of components</i>	20	42	58	137	270

Source: Author's elaboration

ANALYSES IN PROGRESS

The findings of the statistical analyses (*currently in progress*) combined with the qualitative understanding of the case will increase our understanding of which mechanisms are more likely to drive the network dynamics in the case of emerging technologies. These include evidence of the extent to which the likelihood of forming, maintaining, or terminating ties among actors (authors or affiliations) is affected by actors' covariates such as types of organisations (e.g. university, hospital, firms, governmental departments), diversity or specialisation of the research undertaken (e.g. variety of topics), and status (e.g. actors producing highly cited research) as well as how this changes over the emergence. The understating of these dynamics has potential to provide important inputs for policymaking process in the case of emerging technologies.

REFERENCES

- Ahuja, G., Soda, G., & Zaheer, A. (2012). Introduction to the special issue: The genesis and dynamics of organizational networks. *Organization Science*, 23(2), 434–448. doi:10.1287/orsc.1110.0695
- Blume, S. (1992). *Insight and Industry: on the Dynamics of Technological Change in Medicine*. Cambridge, MA: MIT Press.
- Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, 92(5), 1170–1182. doi:10.1086/228631
- Brambilla, D., Luciani, P., & Leroux, J.-C. (2014). Breakthrough discoveries in drug delivery technologies: the next 30 years. *Journal of Controlled Release* □: Official Journal of the
Controlled Release Society, 190, 9–14. doi:10.1016/j.jconrel.2014.03.056

- Burt, R. S. (1992). *Structural Holes: Social Structure of Competition*. London: Harvard Business Press.
- Chesbrough, H. (2003). *Open Innovation: The new Imperative for Creating and Profiting from Technology*. Boston: Harvard Business School Press.
- Collingridge, D. (1980). *The Social Control of Technology*. Milton Keynes: Open University Press.
- Consoli, D., Mina, A., Nelson, R. R., & Ramlogan, R. (2016). *Medical Innovation, Science, Technology and Practice*. London, UK: Routledge.
- Geels, F. W. (2002). Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study. *Research Policy*, 31(8-9), 1257–1274. doi:10.1016/S0048-7333(02)00062-8
- Granovetter, M. S. (1983). The Strength of Weak Ties: A Network Theory Revisited. *Sociological Theory*, 1, 201–233. Retrieved from <http://www.jstor.org/stable/202051>
- Gulati, R., & Gargiulo, M. (1999). Where Do Interorganizational Networks Come From? *American Journal of Sociology*, 104(5), 1439–1438. doi:10.1086/210179
- Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., & Morris, M. (2008). statnet: Software Tools for the Representation, Visualization, Analysis and Simulation of Network Data. *Journal of Statistical Software*, 24(1), 1–11. doi:10.18637/jss.v024.i01
- Koutsonanos, D. G., Vassilieva, E. V, Stavropoulou, A., Zarnitsyn, V. G., Esser, E. S., Taherbhai, M. T., ... Skountzou, I. (2012). Delivery of subunit influenza vaccine to skin with microneedles improves immunogenicity and long-lived protection. *Scientific Reports*, 2, 357. doi:10.1038/srep00357
- Krivitsky, P. N., & Handcock, M. S. (2014). A separable model for dynamic networks. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 76(1), 29–46. doi:10.1111/rssb.12014
- Padgett, J. F., & Powell, W. W. (2012). *The Emergence of Organizations and Markets*. Princeton, New Jersey: Princeton University Press.
- Petersen, A. M., Rotolo, D., & Leydesdorff, L. (2016). A triple helix model of medical innovation: Supply, demand, and technological capabilities in terms of Medical Subject Headings. *Research Policy*, 45(3), 666–681. doi:10.1016/j.respol.2015.12.004
- Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social Networks*, 29(2), 173–191. doi:10.1016/j.socnet.2006.08.002
- Rotolo, D., Hicks, D., & Martin, B. R. (2015). What is an emerging technology? *Research Policy*, 44(10), 1827–1843. doi:10.1016/j.respol.2015.06.006
- Stirling, A. (2009). *Direction, distribution and diversity! Pluralising progress in innovation, sustainability and development*.
- van Lente, H., & Rip, A. (1998). The rise of membrane technology. From rethorics to social reality. *Social Studies of Science*, 28(2), 221–254.



Using network centrality measures to improve national journal classification lists¹

Alesia Zuccala^{*}, Nicolas Robinson-Garcia^{**}, Rafael Repiso^{***} and Daniel Torres-Salinas^{****}

^{*}a.zuccala@hum.ku.dk

Royal School of Library and Information Science, University of Copenhagen, Birketinget 6, Copenhagen S, DK-2300 (Denmark)

^{**}elrobin@ingenio.upv.es

INGENIO (CSIC-UPV), Universidad Politécnica de Valencia (Spain)

^{***}rafael.repiso@gmail.com

EC3metrics spin-off, Universidad de Granada (Spain)
Universidad Internacional de la Rioja (Spain)

^{****}torressalinas@gmail.com

EC3metrics spin-off, Universidad de Granada (Spain)

INTRODUCTION

Research productivity for the scholar is often evaluated on the basis of his/her journal articles; however, specific journals are said to possess a higher measure of impact than others (Archambault & Larivière, 2009; Garfield, 2006; Glänzel & Moed, 2002). When a scholar decides where to publish, (s)he might consider a journal's impact factor. Although Garfield (1973) claimed that citation counts to individual articles will determine the impact factor of a journal, newer evidence points to the contrary: a journal with a high impact factor can also influence an article's readership and subsequent citation rates (Larivière & Gingras, 2010). This "chicken-and-egg" dispute (i.e., *citations first or impact first?*) can be tested, but can still have negative consequences for how journals are selected, rated, listed, and used by policy-makers for developing measures of scholarly performance. For instance, in countries like Denmark and Spain classified journal lists are now being produced and used in the calculation of nationwide performance indicators. As a result, Danish and Spanish scholars are advised to contribute to journals of high "authority" (as in the former) or those within a high class (as in the latter). This can create a few problems.

First, a classification system that is designed to prize older, more established journals is problematic if it fails to acknowledge the role of the new journal. Scholarly research fields escalate and decline over time, and when a new area intensifies, sometimes a specialty journal is created. Data extracted from the Ulrich's periodical database for the period of 1900 to 1999 indicate "compound annual growth rates" for serials and has been used to suggest that "an increase of about 100 refereed papers per year world-wide, results in the launch of a new journal" (Mabe, 2001, p.159). Socio-political climates can further influence these growth rates, yet when a ranked list of journals is generated, the newer journal will inevitably start at a lower position. A scholar may then question or re-think his/her publication strategy. This type of decision making interferes unnecessarily with the natural flow of the learned society. According to Mabe (2001) the 'learned society' is essentially "a self-organizing information

¹ Nicolas Robinson-Garcia is supported by a Juan de la Cierva-Formación Fellowship granted by the Spanish Ministry of Economy and Competitiveness.

system that reflects the growth and specialization of knowledge” (p. 161). The role of a newly created journal is to function as a formal communication outlet where a gap is noticed within this natural, self-organizing system.

The second problem rests with how journal lists are established and revised. Journals in emerging or peripheral fields might fail to make the list in the first place, while others are placed at a lower level or class. Here we will focus on two systems in particular: the ‘authority list’ related to the Danish bibliometric performance (BFI) system and the Integral Classification of Scientific Journals (known by its Spanish acronym, CIRC) for categorizing journals in Spain (Torres-Salinas, et al., 2010). The two systems differ because the first is based on peer-based judgements, while the latter is based on journal metrics and the presence of journal titles in international databases. The aim of this paper is to analyse the potential use of network centrality measures to identify possible mismatches of journal categories.

The paper is structured as follows. First, we describe the Danish and Spanish classification/rating systems. Second, we test a method for assessing and re-classifying journals in these systems, based on two complementary research techniques: 1) journal co-citation analysis and 2) social network centrality measures. These combined methods have previously been used to assess journals (e.g., Ni et al., 2011; Leydesdorff, 2007; McCain, 1991a, b), yet seem to be overlooked in this case as an informative policy-making tool.

TWO APPROACHES: DANISH AUTHORITY VERSUS SPANISH METRICS

Danish Authority List:

In 2009, Denmark developed an authority list of publications, and since this date, this list has been prepared and audited annually by over 350 researchers, across 68 ‘assigned’ disciplines. Journals, book publishers and conference proceedings that pass the auditing process are categorized by the Danish academics as being either a ‘level 1’ outlet (normal) or a ‘level 2’ ‘prestigious’ outlet. According to the Danish bibliometric point system, known generally as the “BFI”, publishing in a level 2 journal leads to a performance point of 3.0 while publishing in a level 1 journal earns a lower point of 1.0. The level 2 journal is expected to be that which covers a maximum of 20 % of world production of articles in the discipline to which it is assigned. Monographs and chaptered volumes also receive points, but we will not elaborate on these details, as they are not relevant to the scope of this study. The important aspect of the BFI system is that at the end of each year, cumulated points are used to determine how much of the Danish government’s basic research funding (25% of the full allotment) is to be re-distributed amongst all universities (see Pedersen, 2010).

Spanish CIRC Classification:

In 2010, a group of Spanish bibliometric experts proposed a categorization of scientific journals for the Social Sciences and Humanities (Torres-Salinas et al., 2010). The classification aims at synthesizing the criteria of Spanish funding agencies for assessing journals from these areas and it is based on their inclusion and rank in a heterogeneous variety of tools and databases (Web of Science, Scopus, ERIH, etc.). Paradoxically, this classification has recently been included as a criterion in the Spanish performance-based evaluation system (Torres-Salinas & Repiso, 2016). CIRC classifies journals into five classes (A+, A, B, C and D). Journals are classified according to their compliance to certain criteria which are based on their inclusion in international databases and their Journal Impact Factor. It differentiates

between Social Sciences and Humanities, as national evaluation standards also differ. Hence a journal may be categorized as B for Social Sciences and A for Humanities. Table 1 presents the CIRC classification criteria.

Table 1. Spanish CIRC criteria for classifying Social Sciences & Humanities journals.

	Social Sciences	Humanities
A+	<ul style="list-style-type: none"> Journals included in the first quartile in the JCR Social Sciences Edition according to their Impact Factor. 	<ul style="list-style-type: none"> Journals indexed in the A&HCI from Thomson Reuters and also positioned in the first quartile in Scopus according to their Impact per Publication (IPP) score.²
A	<ul style="list-style-type: none"> Journals indexed in the SSCI or A&HCI, excluding those indexed in the fourth quartile of the JCR Social Sciences Edition according to their Impact Factor. Journals indexed in Scopus and positioned in the first quartile according to their IPP. 	
B	<ul style="list-style-type: none"> Journals included in the fourth quartile in the JCR Social Sciences Edition according to their Impact Factor. 	<ul style="list-style-type: none"> Humanities journals indexed in ERIH Plus (European Reference Index for Humanities).
	<ul style="list-style-type: none"> Journals indexed in Scopus in the second, third and fourth quartile according to their IPP score (excluding journals with IPP = 0). Spanish journals with a quality label recognized by the Spanish Foundation for Science and Technology (FECYT). 	
C	<ul style="list-style-type: none"> Journals indexed in Scopus with an IPP = 0. Social Sciences journals indexed in ERIH Plus. Journals indexed in the Regional Information System for Scientific Journals in Latin America, the Caribbean, Spain and Portugal catalogue (LATINDEX). 	
D	<ul style="list-style-type: none"> Journals included in the LATINDEX directory but not in its catalogue. 	

METHODS

Here we will compare the Danish and Spanish systems according to how each class of journal – i.e., level 1 and 2, or A+, A, B, C, D - ‘fits’ within a co-citation network. Our objective is to acquire information about the journal’s network centrality within a specific field. The chosen field is Library and Information Science (LIS) and the method of data collection is as follows:

- A data sample (n=3,831 research articles) was extracted from all core indices of the Web of Science (WoS) for the publication year of 2015, and from the Subject Category: “Information Science and Library Science”.
- The sample articles (n=3,831) were submitted to the VOSViewer mapping algorithm (Van Eck & Waltman, 2010) and used to produce a journal co-citation analysis based on a minimum citation threshold per journal set to 111 citations.
- A final co-citation network of 151 nodes was produced in VOSViewer (see Figure 1).
- A Pajek (*.net) file was then extracted from VOSViewer and used as input to the social network analysis and mapping tool, Netdraw (Borgatti, 2002).

² Information obtained from <http://www.journalmetrics.com>

- In Netdraw, a selection of node centrality measures, including *degree*, *betweenness*, *closeness*, and *eigenvector measures*, were calculated for each of the 151 nodes in the LIS journal co-citation network. Our research focuses mainly on the *eigenvector* and *betweenness* measures. *Eigenvector* centrality characterizes the global centrality of a node in a network and it is the most interesting indicator for our study, along with *betweenness* which indicates where a node possesses the shortest path between other node pairs, and shows the least correlation with the rest of the centrality indicators (Valente et al., 2008).

Figure 1. Journal co-citation network (n=151 nodes) from the WoS “Information and Library Science” category and each journal’s Danish versus Spanish classifications.

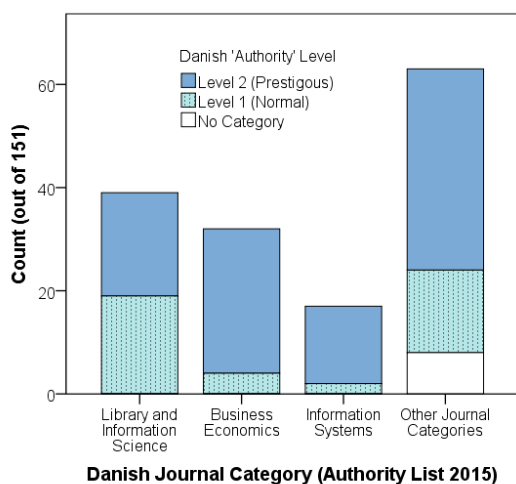
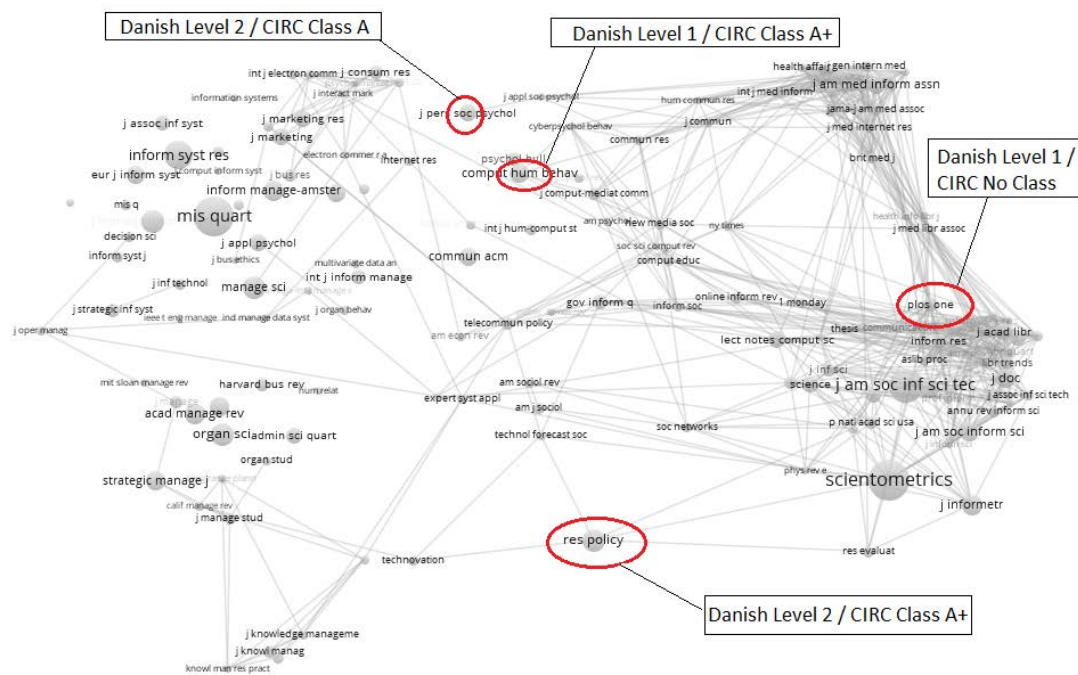


Figure 2

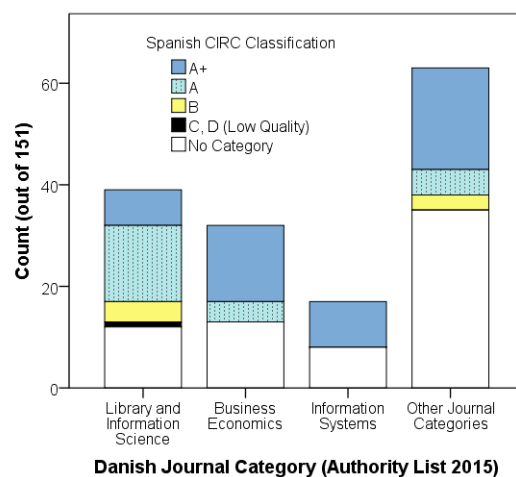


Figure 3

RESULTS

Journal authority versus network centrality:

Note that in Figure 1, we show the natural co-citation patterns of 151 journal titles from Web of Science journal category of “Information Science and Library Science”. Each journal in the co-citation network possesses either a central or peripheral role, or plays an ‘in-between’ role as a bridge between topics. What we expect from a co-citation network of approximately 100-150 journals in a field is that new clusters will form over time; clusters that might lead to the creation of a new journal, or a central position for an existing one.

Figures 2, and 3, above indicate that the journal network has grown out of useful contributions from various fields. In both figures we use field categories as defined in the Danish authority list, but compare the different journal classifications from the Danish Level 1 system versus the CIRC A+, A, B, C, D system. 58% of the co-cited journals in the network are level 2 journals, from the fields of *LIS* (26%), *business economics* (11%) or *information systems* (21%). Less than half of the journals are from other ‘related’ fields (e.g., *computer science, public health, science studies, media & communication, political science, medicine, psychology*).

Figures 4 and 5. Eigenvector centrality values for journals in the 2015 LIS co-citation network: Danish “authority” levels versus Spanish CIRC classifications.

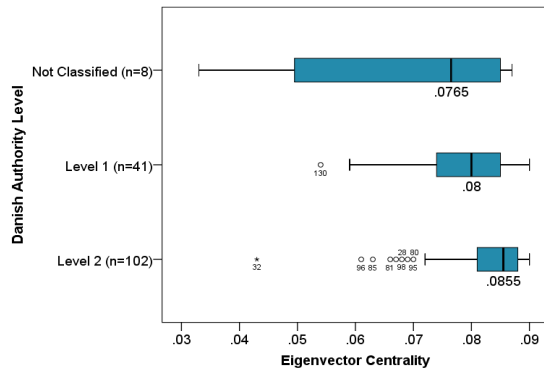


Figure 4

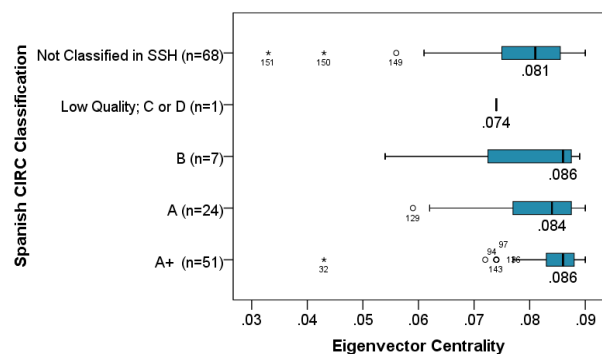


Figure 5.

The boxplots in Figures 4 and 5, show that the median *eigenvector* centrality values for journals classified by the Danish ‘authority’ system at level 1 or 2 differ slightly (.08 and .0855), while the A+, A and B median values in the Spanish system barely differ at all (.086, .084, and .086). Note from Figure 4 that some journals in the third and fourth quartiles of the level 1 boxplot have *eigenvector* values that are just as high as those at or above the median value in the level 2 boxplot. *ASLIB Proceedings* is one example of a journal that has an *eigenvector* of $n=.09$, which is higher than the level 2 boxplot median (.0855). While it has been classified by the Danish system as a level 1 journal, it may have potential to be re-classified at some point to level 2. One concern; however, is that it has had a name change to *ASLIB Journal of Information Management*; hence this needs to be accounted for in a repeated network analysis.

In Figure 5, the boxplot representing all journals with a B rating (Spanish CIRC) is skewed to the left. This indicates that more observations fall below the median, yet there are still a few B journals (above the comparative A .084 median) that play as much a central role in a network as an A or A+ journal. For example, the journal *Information Research* is classified at level 2 in Denmark, but for the Spanish this is a B journal.

Journals classified in both the Danish level 2 and Spanish A+ categories with the highest *eigenvector* centrality measures (.09) include: *Journal of the American Society for Information Science and Technology* (now *Journal of the Association for Information Science and Technology*), the *International Journal of Information Management* and *Scientometrics*. Journals classified in both the Danish level 1 and Spanish B categories with low *eigenvector* centrality measures include: *D-lib Magazine* (.06) and *Reference User Services Quarterly* (.05). The multiple outliers visible in the level 2 boxplot (see Figure 2) represent journals that play a less central role in library and information science, but have a ‘prestigious’ standing in other related fields (e.g., *Journal of Personality and Social Psychology*; *Journal of Marketing Research*).

Figures 6 and 7. Betweenness centrality value for journals in the 2015 LIS co-citation network: Danish “authority” levels versus Spanish CIRC classifications.

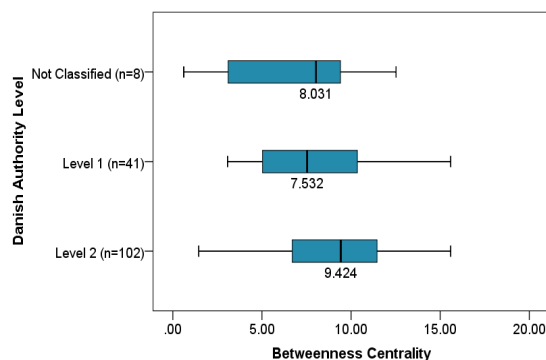


Figure 6

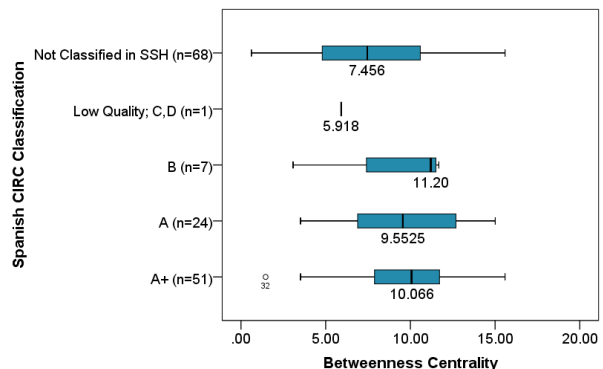


Figure 7

Figures 6 and 7 present boxplot distributions for the Danish and Spanish journal classes based on their *betweenness* measures in the network. In Figure 6, the boxplot for journals classified by the Danish system at level 1 is skewed to the right. This indicates that a higher measure of *betweenness* is observed more often for journals in this class than it is for ‘prestigious’ level 2 journals. Although many level 2 journals associated with outlier fields (e.g., psychology, medicine, economics) also have a high *betweenness* value. Note also from Figure 7 that the B journals classified by the Spanish system also tend to show a higher median value of *betweenness* than those from the A+ or A class.

In Table 2, below, the Danish and Spanish journal classification systems are compared again, and this time the level 1, 2, A+, A, B, C/D journals are distributed by quartiles according to their *eigenvector* centrality value. The results in table 2 may be examined in two different ways. For instance, we can focus on the journals that are considered ‘prestigious’ by both classification systems (level 2 in the Danish List, and A+ or A in the Spanish CIRC classification). 39.2% of the level 2 journals in the Danish list are also included in the top

25% according to their *eigenvector*. A slightly higher value can be observed for the Spanish list (33.3%) and even higher if we only focus on the A⁺ journals (49.0%). If we take an opposite view and examine the distribution of Q1 journals according to their *eigenvector* value, we observe that the highest share of these journals are categorised as prestigious (83.3% for the Danish class, 68.8% for the Spanish CIRC).

Table 2. Grouping of journals based on their Danish authority level and Spanish CIRC classification and their eigenquartile measures.

Danish Authority List					
	Q1	Q2	Q3	Q4	Total
2	40	22	27	13	102
1	7	4	14	16	41
0	1	2	1	4	8
Spanish CIRC Classification					
A ⁺	25	11	9	6	51
A	8	4	7	5	24
B	2	2	1	2	7
C/D	0	0	0	1	1
Not included*	13	11	25	19	68
Total	48	28	42	33	151

* These journals may not be included because a) they are not Social Sciences journals, or b) they have simply not been reported and are missing.

The ‘evolving’ journal

In this part of our study we show how network centrality measures may be used as a support tool for re-classifying journals, particularly for those assigned to lists like the Danish and Spanish systems. Earlier we explained that the development of such lists can be problematic, because they might encourage scholars to publish in certain journals for the wrong reasons, or they can be too rigid if particular journals are kept a specific ‘level’ or class year after year. Our focal point for this analysis is the *Journal of Informetrics*, a relatively a young journal (featured in Figure 1), which was introduced in 2007, and has, within a short period of time, achieved a ‘central’ position in the field of LIS. This journal has been rated highly in both the 2015 versions of the Danish and Spanish journal classifications (i.e., level 2 and A⁺ respectively).

With the *Journal of Informetrics* we have chosen to observe changes to its *eigenvector* centrality over time, alongside the *eigenvectors* of two more journals, *ASLIB Proceedings* and the *Annual Review of Information Science and Technology (ARIST)*. To determine its evolution in the LIS field we have re-iterated similar journal co-citation networks for the earlier years of 2007 and 2011. The two new networks contain the top co-cited 151 nodes, like the 2015 map shown in Figure 1, and each was developed according to the same method as Figure 1. *ASLIB Proceedings* is currently categorized as a level 1 journal in the Danish authority list and A in the Spanish CIRC classification for social sciences. It is positioned in the second quartile according to its *eigenvector* value and serves as one example of a journal that could potentially be re-classified to level 2. *ARIST*, on the other hand is a level 2 journal that has been terminated as of 2011, yet in the year 2015 it was mistakenly classified as a level 1 (Danish) and A⁺ journal (Spanish CIRC).

Figures 8 and 9 present the *eigenvector and betweenness* values for these three journals for the years 2007, 2011 and 2015. The *eigenvector* value shows the global centrality of a journal in a network, thus *ASLIB* but especially *ARIST* were core to the field in 2007. However, *Journal of Informetrics* has an incremental role. For all three journals, we see a convergence over time. The *eigenvector* value of *ARIST* decreases slightly, *ASLIB Proceedings* remains stable and the *Journal of Informetrics* increases.

Figures 8 and 9. *Eigenvector and betweenness* centrality measures of *ARIST*, *ASLIB* and *Journal of Informetrics* in three different periods within the LIS co-citation journal network

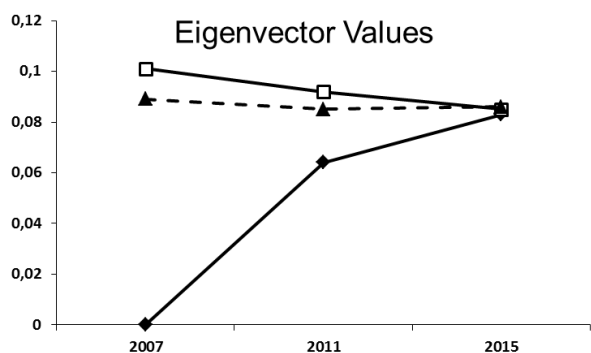


Figure 8.

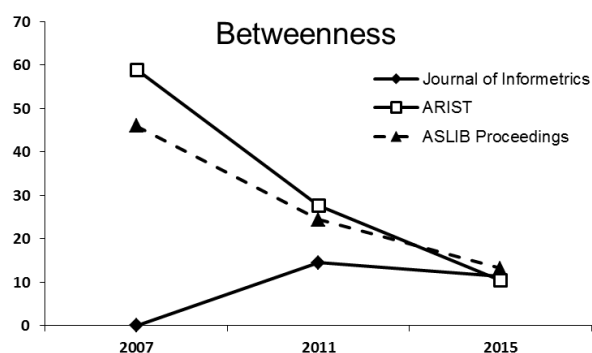


Figure 9.

In Figure 9, *betweenness* values show the local position of nodes in the network, and we are interested in this measure, because it can be used to identify journals that play a ‘brokering’ role (i.e., a link between topics). One might expect a new journal to play this role in its early stages, and indeed this is partially what happens with the *Journal of Informetrics* between 2007 and 2011. After 2011, its role as broker decreases slightly, but it is also within the same period (2011-2015) when it achieves a higher *eigenvector* centrality, and becomes more central to LIS. An overall examination of each journal’s changing *betweenness* measure shows that both *ARIST* and *ASLIB Proceedings* decrease in value, with *ARIST* showing the most dramatic decrease, while *Journal of Informetrics* increases slightly from 2007 to 2011, and then stabilizes.

CONCLUSIONS AND FURTHER RESEARCH

Thus far, both the Danish and Spanish national performance systems have relied either on traditional journal indicators, the presence of journals in international databases, or academic selection committees for developing journal classification lists. Notwithstanding problems associated with journal lists in the first place, this study shows that co-citation network centrality measures might be useful, particularly as a complementary policy tool. Here we conclude with a few policy-related recommendations and suggestions for further research.

While co-citation networks and their centrality measures are not sufficient for establishing the ‘prestige’ of a journal, they can still be used for making adjustments to a journal list. To maintain a list that is reliable, it will definitely require periodic revisions. The journal in question could be a new one, an older one, or one that has ceased to be active; thus a policy might be implemented whereby its centrality measure is observed across five-year periods

within its natural co-citation network. This measure may determine whether or not the journal should be a) introduced at level 1, b) stay in its current level, c) re-assigned as an ‘evolved’ level 2 journal, or d) removed from a list entirely (as with ARIST). In the case of a multidisciplinary journal, such as *Plos One* or *Nature*, the centrality measure will be different if it appears in different networks; thus consistent, relatively high centrality measures in multiple networks could be used to decide its classification.

Since we have focused on the *Journal of Informetrics*, it is important to note that the Danish ‘authority’ list was not established at the time this journal was first published in 2007. We do not know if it would have been added to the Danish ‘authority’ list as a level 1 journal in 2007 before it was ‘promoted’ to level 2. However, because it has experienced a rapid periphery-to-core transition within the LIS field, it seems to have earned its present level 2 class. Additional measures, factors, or dimensions may have also contributed to its growth (see Haustein, 2007), but its network eigenvector centrality is still useful as a complementary ‘objective’ measure.

Last but not least, we need to consider future research. Overall there seems to be a general bias with older journals assigned to a higher class. For example, the founding year for journals listed as A+ in CIRC, is between the mid-1970s up to the mid-1980s, and as the establishing year of the journal gets higher (after 1990) the average class gets lower (as in the case of C journals for Humanities). The evolution of the *Journal of Informetrics* could be exceptional. Many new journals might not show a similar periphery-to-core evolution in a co-citation network. It will be useful therefore to compare this journal’s centrality shifts to other newer journals established at the same time, or to new journals from other fields.

REFERENCES

- Archambault, E & Larivière, V. (2009). History of the journal impact factor: Contingencies and consequences. *Scientometrics*, 79(3), 635–649.
- Borgatti, S.P. (2002). *NetDraw software for network visualization*. Analytic Technologies: Lexington, KY
- Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, 27(1), 55-71.
- Freeman, L. C. (1979). Centrality in social networks conceptual clarification. *Social Networks*, 1(3), 215-239.
- González-Pereira B., Guerrero-Boteb, V.P., & Moya-Anegón, F. (2010). A new approach to the metric of journals’ scientific prestige: The SJR indicator. *Journal of Informetrics*, 4(3), 379-391.
- Garfield, E. (1955). Citation indexes for science. A new dimension in documentation through association of ideas. *Science*, 122, 108-111.
- Garfield, E. (1973). Citation impact depends upon the paper, not the journal! Don’t count on citation by association. *Current Contents*, 22, 5–6.

Garfield, E. (2006). The history and meaning of the Journal Impact Factor. *JAMA*, 295(1), 90-93. doi:10.1001/jama.295.1.90.

Glanzel, W. & Moed, H. (2002). Journal impact measures in bibliometric research. *Scientometrics*, 53(2), 171-193.

Haustein, S. (2012). *Multidimensional journal evaluation. Analyzing scientific periodicals beyond the impact factor*. Berlin: Walter de Gruyter GmbH & Co.

Larivière, V. & Gingras, Y. (2010). The Impact Factor's Matthew Effect: a natural experiment in bibliometrics. *Journal of the American Society for Information Science and Technology*, 61(2), 424-42.

Leydesdorff, L. (2007). Betweenness centrality as an indicator of the interdisciplinarity of scientific journals. *Journal of the American Society for Information Science and Technology*, 58(9), 1303-1319.

McCain, K. W. (1991a). Core journal networks and cocitation maps: new bibliometric tools for serials research and management. *Library Quarterly*, 61(3), 311-336.

McCain, K. W. (1991b). Mapping economics through the journal literature: an experiment in journal co-citation analysis. *Journal of the American Society for Information Science*, 42(4), 290-296.

Moed, H.F. (2011). The source normalized impact per paper is a valid and sophisticated indicator of journal citation impact. *Journal of the American Society for Information Science and Technology*, 62(1), 211-213.

Ni, C., Sugimoto, C., Jiang, J. (2011). Degree, closeness, and betweenness: application of group centrality measurements to explore macro-disciplinary evolution diachronically. In E. Noyons, P. Ngulube, & J. Leta, *Proceedings of the 13th International Conference of the International Society for Scientometrics and Informetrics*, Vol. 1. Durban, South Africa, 04-07 July.

Pedersen, C. V. (2010). The Danish Bibliometric Research Indicator-BFI Research publications, research assessment, university funding. Sciecom Info-Nordic-Baltic Forum for Scientific Communication, 4, 1-4. Accessed on February 18, 2016 from <http://journals.lub.lu.se/index.php/sciecominfo/article/view/4757/4318>.

Torres-Salinas, D., Bordons, M., Giménez-Toledo, E., Delgado López-Cózar, E., Jiménez-Contreras, E. & Sanz-Casado, E. (2010). Clasificación integrada de revistas científicas (CIRC): propuesta de categorización de las revistas en ciencias sociales y humanas. *El profesional de la información*, 19(6), 675-683.

Torres-Salinas, D. & Repiso, R. (2016). Clasificación CIRC 2016: transitando un viejo mundo que se muere. *Anuario ThinkEPI*, <https://listserv.rediris.es/cgi-bin/wa?A2=IWETEL;f881ce57.1602b>

Valente, T.W., Coronges, K., Lakon, C., & Costenbader, E. (2008). How Correlated Are Network Centrality Measures? *Connections*, 28(1), 16–26.

Van Eck, N.J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538.

Wasserman, S. & Faust, K. (1994). *Social network analysis: methods and applications*. New York: Cambridge University Press.



Bridging centrality: A new indicator to measure the positioning of actors in R&D networks¹

Thomas Scherngell*, Iris Wanzenböck* and Laurent Bergé**

*thomas.scherngell@ait.ac.at, iris.wanzenboeck@ait.ac.at
Innovation Systems Department, AIT Austrian Institute of Technology, Vienna, 1220 (Austria)

** laurent.Berge@parisdescartes.fr
MAP5 (UMR CNRS 8145), Université Paris Descartes & Sorbonne, Paris Cité (France)

ABSTRACT

In the recent past, we can observe growing interest in the STI community in the notion of *positioning indicators*, shifting emphasis to actors in the innovation process and their R&D inter-linkages with other actors. In relation to this, we suggest a new approach for assessing the positioning of actors relying on the notion of bridging centrality (BC). Based on the concept of bridging paths, i.e. a set of two links connecting three actors across three different aggregate nodes (e.g. organisations, regions or countries), we argue that triangulation in networks is a key issue for knowledge recombinations and the extension of an actor's knowledge base. As bridges are most often not empirically observable at the individual level of research teams, we propose an approximated BC measure that provides a flexible framework for dealing with the aggregation problem in positioning actors. Hereby, BC is viewed as a function of an aggregate node's (i) participation intensity in the network, (ii) its openness to other nodes (i.e. the relative outward orientation of network links), and (iii) the diversification of links to other nodes. In doing so, we provide an integrative perspective that enables us to achieve a better understanding of the positioning of certain actors in R&D networks. An illustrative example on the co-patent network of European regions demonstrates the performance and usefulness of our BC measure for networks constructed at the aggregated level, i.e. regions in our example. A region's outward orientation and the diversification of its network links moderates the influence of regional scale on network centrality. This is a major strength of the measure, and it paves the way for future studies to examine the role of certain aggregate node's, and, by this, contributes to the debate on positioning indicators in the STI context.

INTRODUCTION

Over the past decade, we have observed considerable progress in the advancement and application of Science, Technology and Innovation (STI) indicators (see, e.g., OECD 2005). In this context, the notion of *positioning indicators* has come into fairly wide use in the STI community. It originates from considerations on new requirements imposed to the production of STI indicators in terms of their adaption from classical input-output to a positioning indicators framework, focusing on flows and linkages between research actors in the innovation system (Lepori 2008). These linkages materialize in form of more formalised collaborations in R&D, such as joint R&D projects (see, e.g., Scherngell and Barber 2009 and 2011, Scherngell and Lata 2013), joint publication activities (see, e.g. Glänzel and Schubert 2004), and researchers mobility (see e.g. Edler et al. 2011). Similarly, they may appear as informal knowledge flows - often referred to as disembodied knowledge spillovers (see, e.g.,

¹ This work has been supported by the Austrian National Bank (ÖNB), Jubiläumsfonds, project No. 16301

Fischer et al. 2006). Despite the fact that a number of works have provided evidence on the crucial importance of R&D linkages, most STI indicators remain rooted in a classical, linear conception of the innovation process.

In this study, we employ a network perspective on R&D linkages. A *R&D network* comprises a set of nodes representing knowledge producing actors inter-linked via edges representing knowledge flows. In Social Network Analysis (SNA), the *positioning* of actors is usually described by the concept of *centrality*, trying to capture a certain function and/or role a node, i.e. actor, takes by means of its inter-linking to other nodes (Borgatti 2005). Thus, the concept of *centrality* can be very well related to the notion of *positioning indicators*. Up to now, only a few STI studies have started to utilize the centrality concept to capture the positioning of actors in innovation systems, usually employing most basic analytical concepts, such as degree centrality or betweenness centrality (see, e.g., Wanzenböck et al. 2014 and 2015, Heller-Schuh et al. 2011). However, they somehow neglect conceptual problems that arise if networks are defined at the aggregate level of (large) organisations or even regions and countries, rather than the most relevant level in terms of R&D linkages, usually a research team within an organisation. Furthermore, the used centrality measures are not conceptually adapted to the STI context, such as the incorporation of theoretical considerations on the diversification of links or the relative outward orientation of R&D networks.

Thus, we propose a measurement approach that specifically adapts common centrality measures to STI relevant theoretical considerations, and provides a flexible framework to overcome problems related to node aggregation of the network. We shift attention to the notion of *Bridging Centrality* (BC), based on the concept of inter-nodal *bridging paths*, denoting an indirect connection between two nodes via a third ‘bridging node’. However, since bridges are usually difficult to be observed at the level of individual researchers, our proposed BC measure shows properties that allow us to estimate the centrality of an aggregate node based on the underlying micro-structure of the network. The objective of this study is to introduce the conception and formal derivation of BC, and demonstrate its interpretative power by an illustrative example. Our approximation of the number of bridging paths of an aggregate node is given by the product of three components, that is, first, a node’s *participation intensity* in the network, i.e. its number of links, ii) a node’s *relative outward orientation*, i.e. the ratio of node-internal (loops) vs. node external links, and (iii) the *diversification* of links across other nodes in the network. Emanating from our conceptual discussion, we provide a formal proof on how our measure decomposed into these three components converges mathematically to a node’s expected number of bridges. Since all three components are relevant for STI studies on its own, the measure shows high interpretative power and, by this, significantly enriches our toolset of positioning indicators, not only in terms of a more appropriate centrality measure.

The remainder of this study is structured as follows: The next section introduces the network notation to elaborate in some more detail on the notion of bridges. Afterwards we outline the formal definition of our BC measure based on the three components *participation intensity*, *relative outward orientation* and *diversification*. It shows how we conceptually perceive the number of bridges of a node starting from these three components, before we provide a formal proof that our measurement approach mathematically corresponds to the expected number of bridges of a node. Then we shift attention to an illustrative example, where we apply our measure to the European co-patent network observed at the level of NUTS-2 regions, and compare the results with conventional centrality measures as well as with respect to the three

BC components. The final section closes with some concluding remarks and ideas for a future research agenda.

Network definition and the concept of bridges

In social sciences, analytical strategies employed to deal with the divide between individualistic and holistic approaches for describing social systems are referred to as multilevel analysis (see, e.g., Lazega and Snijders 2015). In traditional sociological literature, this is aptly described as the phenomenon of *ecological fallacy*, pointing to logical failures in the inference of statistical data observed at an aggregated level on the nature and characteristics of individuals (see Robinson 1950). Social Network Analysis (SNA) faces, on the one hand, similar problems when applied to aggregate nodes, in particular in a STI context (see, e.g., Wanzenböck et al. 2014), while on the other hand entails promising potential to overcome such analytical problems (Snijders 2016).

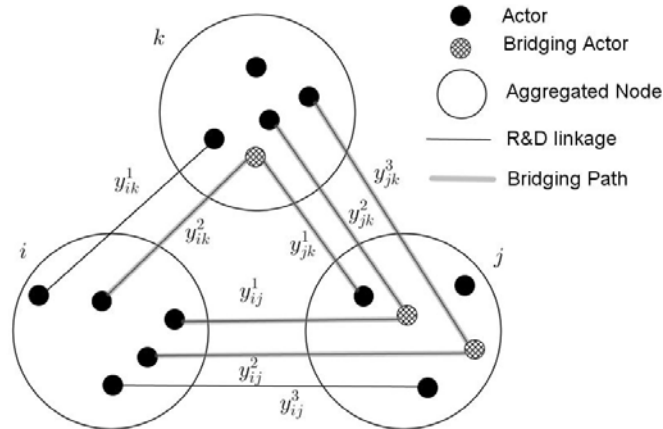
We argue that these aggregation problems prominently occur in the measurement of the positioning of actors in STI studies. Shifting attention to positioning in a network analytic context, we draw on the rich SNA toolbox to evaluate the positioning. The concept of centrality is fundamental in this respect, usually adopted to assign a value to each actor depending on their position within the network (Wasserman and Faust 1994). However, most SNA measures of centrality have been developed for the analysis of social systems, where the nodes of the network are usually identified in terms of individual persons. Accordingly, the original meaning borne by the SNA centrality measures as well as respective interpretations rely on assumptions on the social behaviours of individual persons, and how these persons might influence each other by these behaviours. Centrality measures based on observations at the aggregate level (e.g. organisations, regions or countries) therefore raise important conceptual issues. Most importantly, it implies that every individual actor of an *aggregate node* would homogeneously benefit from the R&D linkages to other nodes, irrespective of who establishes the relations and the strength of these relations.

We propose a flexible analytical approach to address conceptual problems related to unobserved micro-level structures of the observed network. Core in this context is the concept of *bridging path* denoting a form of indirect connection between aggregate nodes. For a formal definition, consider a network observed at the level of aggregate nodes, e.g. organisations, regions or countries, and the connections between the aggregate nodes represent the R&D linkages between their individual actors. This represents a weighted network where we define g_{ij} as the number of R&D linkages (i.e. micro-level links) between aggregate nodes i and j . Further, each micro-level link between two aggregate nodes is denoted by y_{ij}^a , representing the a^{th} link between aggregate nodes' i and j with $a \in \{1, \dots, g_{ij}\}$. A pair of links (y_{ik}^a, y_{jk}^b) forms a bridging path if, and only if, y_{ik}^a and y_{jk}^b are connected to the same actor of aggregate node k . This concept is depicted by Figure 1 exemplified with three aggregated nodes.

The concept of bridges is of particular relevance in a STI context. A high number of bridging paths implies a more open positioning in the network. In contrast to closed and dense network structures, such a bridging position between other nodes can be related to the access to a more diversified knowledge pool. It is assumed that the sources from which the individual actors draw their knowledge will have an impact on their ability to generate innovations, and knowledge flowing through bridging paths is more likely heterogeneous and non-redundant.

Based on these conceptual considerations on bridging paths, we propose a measurement approach for *Bridging Centrality (BC)* in the section that follows.

Figure 1: Illustration of the notion of bridging path



Notes: The figure depicts three bridging paths formed by the following pairs of links: (y_{ik}^2, y_{ik}^1) , (y_{ij}^1, y_{ij}^2) and (y_{ij}^2, y_{ij}^3) . So the aggregate node dyads (j, k) , (i, k) and (i, j) have respectively 0, 2 and 1 bridging paths stemming from aggregate nodes i, j and k .

Three components of Bridging Centrality

Given the parsimonious and effective formal definition of bridging paths, it could be assumed at a first glance that the definition of a formal BC measure is straightforward. Indeed, this is the case in pure mathematical terms as the *true* measure of BC for an aggregate node i would just be the number of bridging actors assigned to i , probably normalised by the total of all bridging actors in the network. However, de facto we are often confronted in social sciences with a well-known problem of finding appropriate empirical observations for the objects under scrutiny. This is particularly critical in STI studies, where we usually focus on large-scale networks such as co-patent, co-publication or project networks. Most often information on links at the level of the individual researchers cannot be traced; even when information is available (as e.g. for authors in publications and/or inventors in patents), the observation for large-scale networks is infeasible due to immense efforts for data cleaning, in particular name standardisation over space and time.

Thus, we propose an alternative measure for BC that approximates the number of bridges of an aggregate node. Drawing on theoretical considerations from various STI studies, we assume that the number of bridging actors of an aggregate node may to a large extent depend on three components: the node's i) *participation intensity*, ii) *relative outward orientation* and iii) *diversification* of network links. We will show that a linear-multiplicative combination of these components formally dissolves to the *expected number of bridges* (see Bergé 2016). At the same time, the three components of BC are highly relevant, each of them having important mechanisms on its own and significant implications on knowledge creation structures.

In our formal description, we denote C_i as the approximated BC for the aggregate i node by

$$C_i = q_i s_i (1 - h_i) \quad (1)$$

where

q_i is the weighted degree of aggregate node i , defined as the total number of links excluding node-internal ones, i.e. $q_i = g_i - g_{ii}$. It refers to the overall *participation intensity* in the network; an aggregate node's size will amplify the probability of yielding more bridges between other nodes.

s_i is the *relative outward orientation* of aggregate node i with $s_i = q_i / g_i$. It reflects the degree of openness of an aggregate node with respect to all established links. Given the focus on bridges, the capacity of an aggregate node to link to other nodes would decrease by a higher number of node-internal links (loops) as it potentially reduces the number of actors connecting different aggregate nodes.

h_i refers to the degree of *diversification* of network links of aggregate node i among other nodes. It is measured by the Herfindahl-Hirschman (HH) index by $h_i = \sum_{j \neq i} (g_{ij} / q_i)^2$. The term $1 - h_i$ varies between 0 and 1, and indicates how an aggregate node's linkages are distributed along its neighbouring nodes in the network. The more the linkages are concentrated, the less the node is central in terms of BC. Concentration reduces the actors' possibility to build bridges among different aggregate nodes and to draw its knowledge from different sources.

An aggregate node's ability to benefit from new ties in the R&D network or exploit external knowledge sources via the links may be determined by all three components together. Outward orientation and higher diversification in particular may help actors belonging to an aggregate node to develop and renew their knowledge base faster, or prevent lock-in situations in certain technologies (see, e.g., Breschi et al. 2015). Hence, our measure to approximate BC features promising opportunities in terms of interpretation.

However, it is not only conceptually attractive, but also mathematically corresponds very well with the *Expected Number of Bridges* (ENB) measure as introduced by (Bergé 2016) using basic random matching assumptions between aggregate nodes². Mathematically, our measure simply collapses to ENB, given by

$$\begin{aligned}
 C_i &= q_i s_i (1 - h_i) = \frac{q_i^2}{g_i} \left(1 - \sum_{j \neq i} (g_{ij} / q_i)^2 \right) = \frac{q_i^2}{g_i} - \frac{1}{g_i} \sum_{j \neq i} g_{ij}^2 \\
 &= \frac{1}{g_i} \sum_{j \neq i} g_{ij} (q_i - g_{ij}) = \frac{1}{g_i} \sum_{j \neq i} (g_{ij} \sum_{k \neq i, j} g_{ik}) \\
 &= \sum_{j \neq i} \sum_{k \neq i, j} \frac{g_{ij} g_{ik}}{g_i}
 \end{aligned} \tag{2}$$

with the latter expression corresponding to the most general form of the expected number of bridges. This can of course be extended to more reasonable assumptions of expected bridges, for instance, by considering the number of actors of aggregate nodes proportional to the

² Note in this context that the random matching process is 'noisier' the larger the aggregate node is, e.g. when nodes are countries

number of R&D linkages (see Bergé 2016 for details). This version of the expected number of bridges corresponds to empirically observed cases quite well.

An application to European cross-region patent networks

In this section, we shift attention to a compact illustration of the BC measure in order to demonstrate its behaviour and interpretative power. We use the example of co-patent network data that is observed at the level of European NUTS-2 regions to represent the aggregate level in our network, and compare the BC with three other commonly used centrality measures, that is the degree, the eigenvector and the betweenness centrality (Wasserman and Faust 1994)³. A co-patent is regarded as a collaboration of at least two inventors issuing a patent grant, providing us a trail of R&D linkages. Respective data are extracted from the REGPAT database (see Maraut et al. 2008) and consist of all patents applied for at the European patent office (EPO) in the period 2006-2010. Our cross-regional co-patenting network is based on a total of 171,451 patents, producing 121,036 inter-regional collaborations linking 250 NUTS-2 regions (see Bergé et al. 2015 for further details on the data)⁴. In this example, the aggregation problem described in Section 2 clearly applies as we are not able to observe co-patent activities at the level of individual inventors⁵.

Table 1 represents the top 30 centralities ordered by the bridging centrality. We focus on commenting the most salient differences. The ranking is clearly dominated by German regions which rank highest for most measures⁶. However, the concentration tendency and high clustering of co-patenting activities at the national level of Germany may point to the fact that economic linkages at the national level prevail. Likely explanations are low cultural barriers as well as lower transaction costs. These factors seem to promote the high regional bridging centrality in German regions⁷.

Further interesting specific cases are, e.g., Île de France (FR10) or Brussels (BE10). FR10 ranks at the 16th position for BC, while being ranked first with respect to its degree centrality. Degree centrality may overstate its position; Although the structure of the collaborations of FR10 with its partnering regions is highly distributed (it has a low HH index of 0.04), this region is characterised by a high number of internal collaborations (the outer share of collaborations is only 44%), and thus, do not provide many bridging paths to the inter-regional R&D network. BE10 ranks below 55th for degree and eigenvector centrality, while for BC it ranks 30th. These conventional centrality measures may underestimate its positioning in the inter-regional co-patent network due to its very high outward orientation

³ The degree is here calculated as the number of unique R&D interactions the agents of a region are involved in. The eigenvector and the betweenness centrality are computed using the package *igraph* available in the statistical software R. Both these two measures are based on the weighted regional co-patent network where the nodes are the regions and where the linkages between any two regions are the number of patents co-invented by agents from these two regions. Due to the nature of the network, we used the weighted version of both the betweenness and the eigenvector centrality.

⁴ Note that the use of different time frames to build the dataset, such as 2004–2006 or 2008–2010, imply no important changes on the results.

⁵ An aggregation to the organisational level would also be inconsistent, as patents are most often assigned to headquarters of companies which often does not reflect to the locus of knowledge creation.

⁶ Note that the performance of German regions is not merely driven by the fact that German NUTS2 regions are usually smaller geographical aggregates than NUTS2 regions in other EU countries, which could drive up their number of inter-regional collaborations at the national level. Indeed, when we redo the analysis taking German regions at the NUTS1 level while keeping other regions at the NUTS2 level, German regions still trust the top of the rankings.

⁷ The national versus international nature of collaborations and its effects on regional network centrality might deserve further attention, and constitute an interesting route for the further development of the regional bridging centrality measure.

(its outer share is 94%) and a highly distributed structure of collaborations (it has a low HH index of 0.07); BE10 is likely to provide many bridging paths to the network.

Table 1: Centrality values of the top 30 regions for the co-patent network (ranks in brackets)

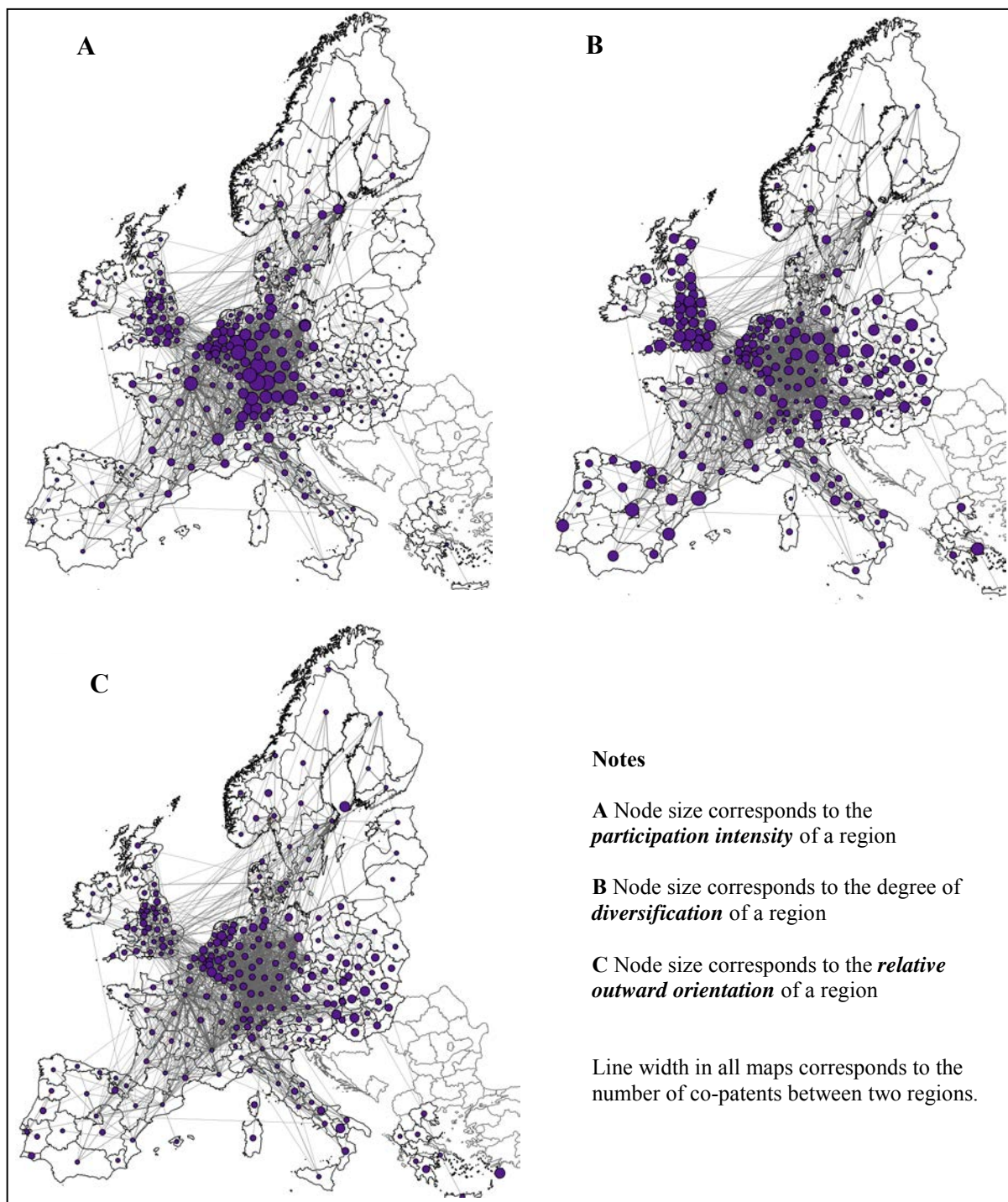
NUTS-2 Code	Region	BC (C_i)	Degree Centrality	Eigenvector Centrality	Betweenness Centrality
DE12	Karlsruhe	1.00 (1)	0.87 (5)	1.00 (1)	0.22 (10)
DE71	Darmstadt	0.93 (2)	0.88 (4)	0.82 (3)	0.45 (4)
DEA1	Düsseldorf	0.84 (3)	0.82 (6)	0.68 (4)	0.22 (9)
DEA2	Köln	0.76 (4)	0.73 (7)	0.63 (6)	0.33 (6)
DEB3	Rhein Hessen-Pfalz	0.73 (5)	0.64 (8)	0.85 (2)	0.13 (16)
DE21	Oberbayern	0.63 (6)	0.96 (2)	0.42 (7)	1.00 (1)
DE11	Stuttgart	0.59 (7)	0.95 (3)	0.64 (5)	0.37 (5)
DE13	Freiburg	0.49 (8)	0.52 (10)	0.34 (9)	0.19 (11)
CH03	NW Switzerland	0.43 (9)	0.41 (14)	0.16 (17)	0.10 (24)
DEA5	Arnsberg	0.42 (10)	0.39 (16)	0.33 (10)	0.06 (45)
DE14	Tübingen	0.40 (11)	0.44 (12)	0.38 (8)	0.07 (36)
DE30	Berlin	0.39 (12)	0.40 (15)	0.22 (14)	0.19 (12)
DEA3	Münster	0.39 (13)	0.31 (20)	0.27 (11)	0.05 (49)
DE25	Mittelfranken	0.37 (14)	0.43 (13)	0.20 (15)	0.11 (20)
CH04	Zurich	0.35 (15)	0.34 (18)	0.12 (21)	0.08 (32)
FR10	Île de France	0.34 (16)	1.00 (1)	0.08 (35)	0.93 (2)
DE27	Schwaben	0.33 (17)	0.31 (21)	0.25 (12)	0.03 (71)
DE40	Brandenburg	0.28 (18)	0.22 (30)	0.15 (18)	0.05 (54)
DE60	Hamburg	0.27 (19)	0.23 (29)	0.09 (28)	0.05 (48)
DE26	Unterfranken	0.27 (20)	0.27 (23)	0.25 (13)	0.10 (23)
FR42	Alsace	0.26 (21)	0.23 (27)	0.13 (19)	0.09 (31)
CH02	Espace Mittelland	0.26 (22)	0.27 (22)	0.08 (30)	0.05 (50)
BE24	Vlaams-Brabant	0.25 (23)	0.20 (34)	0.04 (46)	0.10 (25)
DE92	Hannover	0.24 (24)	0.25 (24)	0.12 (22)	0.05 (53)
FR71	Rhône-Alpes	0.24 (25)	0.57 (9)	0.08 (34)	0.33 (7)
DEB1	Koblenz	0.21 (26)	0.17 (46)	0.18 (16)	0.01 (96)
DE93	Lüneburg	0.21 (27)	0.17 (42)	0.07 (37)	0.02 (79)
CH05	Eastern Switzerland	0.21 (28)	0.19 (36)	0.07 (38)	0.01 (97)
BE21	Prov. Antwerpen	0.20 (29)	0.18 (38)	0.05 (44)	0.09 (28)
BE10	Région de Bruxelles	0.20 (30)	0.14 (59)	0.03 (55)	0.08 (34)

To complement the results from Table 1 in this respect, Table 2 provides a snapshot on top-5 regions including Île de France (FR10) and their respective results across the three components. It becomes obvious that the high rankings of German regions result from the fact that they show both a high participation intensity and openness, i.e. a high absolute as well as relative number of inter-regional co-patents.

Table 2: Ranking of top regions decomposed by three components of BC

Rank	NUTS2	BC	Participation Intensity	Outward Orientation	Diversification
1	DE12	1.00	2333	0.85	0.88
2	DE71	0.85	2050	0.78	0.93
3	DEB3	0.80	1831	0.92	0.84
4	DEA1	0.80	1993	0.80	0.87
5	DEA2	0.78	1866	0.84	0.87
13	FR10	0.34	1382	0.49	0.96

Figure 2: Bridging centrality of regions in the European co-patent network decomposed by three components



In addition, Figure 2 provides an illustrative overview in form of spatial network maps of the European co-patent network, decomposed by the three components. The results are highly interesting, both in terms of illustrating the functioning of BC, as well as in terms of providing insights into the spatial dynamics of European co-patenting. It demonstrates why some regions, such as Ile de France, do not appear on top in terms of BC due to their lower relative outward orientation. Further, commenting on the overall picture, it can be seen the we observe

the classical regions Ile de France as well as regions in Western Germany, the Netherlands and UK to come up with highest participation intensity, while Eastern and Southern European regions show a higher diversification of their links, i.e. actors in that regions seem to be not that selective in the choice of their partners as actors with a high reputation in European core regions. With respect to outward orientation, we also observe high values of some Southern and Eastern European regions. In terms of established co-patent links they seem to be highly open, which could be explained by their reliance on external collaborations and knowledge sources, as well as the lack of internal collaboration structures.

Summary and conclusions

In the STI community, we can observe a lively debate on the notion of positioning indicators, shifting emphasis to actors in the innovation process and their R&D interactions with other actors. However, up to now only few indicators exist which are able to provide a rather comprehensive assessment of an actor's positioning in the innovation system, reflecting structural characteristics of its internal and external R&D linkages. In this study, we suggest a new approach for assessing the positioning of actors in innovation systems relying on the notion of bridging centrality (BC). Based on the concept of bridging paths, i.e. a set of two links connecting three actors across three different aggregate nodes (e.g. organisations, regions or countries), we argue that triangulation in networks is a key issue for knowledge recombination and the extension of an actor's knowledge base.

As bridges are most often not empirically observable at the individual level of research teams, we propose an approximated BC measure that provides a flexible framework for dealing with the aggregation problem in positioning actors. Hereby, BC is viewed as a function of a node's (i) participation intensity in the network, (ii) its openness to other nodes (i.e. the relative outward orientation of network links), and (iii) the diversification of links to other nodes. With these three components – which are both intuitive and computationally simple – we provide an integrative perspective that enables us to achieve a better understanding of the role of certain actors in R&D networks.

An illustrative example on the co-patent network for European regions demonstrates the performance and usefulness of our BC measure for networks constructed at the aggregated level. Despite observing similar patterns in basic statistics like correlations of the centralities, we were able to show striking and interesting differences in the structure of the inter-regional co-patent linkages across regions. A region's outward orientation and the diversification of its network links moderates the influence of regional scale on network centrality. This is a major strength of the measure proposed in this study, and it paves the way for future studies to examine the role of certain aggregate node's, not only regions, but also organisations, in R&D networks, and, by this, contributes to the debate on positioning indicators in the STI context.

Of course, there is room for further improvements of the approach. Indeed, a crucial point for future research is to devote higher emphasis to the specific characteristics of R&D network links and our concept could be used to integrate these aspects. For example, extensions of the bridging centrality could include a focus on the fact that some bridging agents indirectly connect national actors with international ones. By focusing on technology related issues, one could consider bridging agents who indirectly connect actors from one specific technology with others from another technology. Moreover, the measure of bridging centrality is not

limited to the context of R&D but may prove to be useful also for the application in other types of network structures, such as trade flows or economic value chains.

References

- Bergé, L. (2016). Network proximity in the geography of research collaboration. *Papers in Regional Science* [forthcoming].
- Borgatti, S.P. (2005). Centrality and network flow. *Social Networks*, 27(1):55–71.
- Breschi, S. and Lenzi, C. (2015). The Role of External Linkages and Gatekeepers for the Renewal and Expansion of US Cities Knowledge Base, 1990-2004. *Regional Studies*, 49(5): 782–797.
- Burt, R.S. (1992). *Structural holes: The social structure of competition*. Harvard University Press, Cambridge, MA.
- Edler, J., Fier, H. and Grimpe, C. (2011). International scientist mobility and the locus of knowledge and technology transfer. *Research Policy*, 40(6):791–805.
- Fischer, M.M., Scherngell, T. and Jansenberger, E.M. (2006). The geography of knowledge spillovers between high-technology firms in Europe. Evidence from a spatial interaction modelling perspective, *Geographical Analysis* 38(3), 288-309
- Glänzel, W. and Schubert, A. (2004). Analysing scientific networks through co-authorship. In *Handbook of quantitative science and technology research*, pages 257–276. Springer.
- Heller-Schuh, B., Barber M.J., Henriques, L., Paier, M., Pontikakis, D., Scherngell, T., Veltri, G.A. and Weber, M. (2011). *Analysis of Networks in European Framework Programmes (1984-2006)*. Sevilla, European Commission [ISBN: 9789279197147].
- Lazega, E. and Snijders, T. (Ed.) (2015). *Multilevel network analysis for the social sciences*. Springer.
- Lepori, B., Barré, R. and Filliatreau, G., (2008). New perspectives and challenges for the design and production of s&t indicators. *Research Evaluation*, 17(1):33–44.
- Maraut, S., Dernis, H., Webb, C., Spiezia, V. and Guellec. D. (2008). The OECD REGPAT Database. *OECD Science, Technology and Industry Working Papers*.
- OECD (2005) Oslo manual: Guidelines for collecting and interpreting innovation data. *Organisation for Economic Cooperation and Development (OECD), Paris*.
- Robinson, W.S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, 15(3):351–357.
- Scherngell, T. (Ed.) (2013). *The Geography of Networks and R&D Collaborations*. Springer-Physica Verlag, Berlin-Heidelberg-New York.
-

Scherngell, T. and Lata, R. (2013). Towards an integrated European Research Area? Findings from Eigenvector spatially filtered spatial interaction models using European Framework Programme data, *Papers in Regional Science* 92, 555-577.

Scherngell, T. and Barber, M. (2011). Distinct spatial characteristics of industrial and public research collaborations: Evidence from the 5th EU Framework Programme, *The Annals of Regional Science* 46(2), 247-266.

Scherngell, T. and Barber, M. (2009). Spatial interaction modelling of cross-region R&D collaborations. Empirical evidence from the 5th EU Framework Programme, *Papers in Regional Science* 88(3), 531-546.

Snijders, T. (2016). The multiple flavours of multilevel issues for networks. In *Multilevel Network Analysis for the Social Sciences*, pages 15–46. Springer.

Wanzenböck, I., Scherngell, T. and Brenner, T. (2014). Embeddedness of regions in European knowledge networks. A comparative analysis of inter-regional R&D collaborations, co-patents and co- publications, *The Annals of Regional Science* 53, 337-368.

Wanzenböck, I., Scherngell, T. and Lata, R. (2015). Embeddedness of European regions in EU funded R&D networks: A spatial econometric perspective, *Regional Studies* 49(10), 1685-1705.

Wasserman, S. and Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. 8). Cambridge university press.



Network heterogeneity in an undirected network¹

Xiaojun Hu,^{*} Loet Leydesdorff,^{**} and Ronald Rousseau^{***}

^{*} xjhu@zju.edu.cn

Medical Information Center, Zhejiang University School of Medicine, Hangzhou 310058 (China)

^{**} loet@leydesdorff.net

University of Amsterdam, Amsterdam School of Communication Research (ASCoR), PO Box 15793, 1001 NG Amsterdam (The Netherlands)

^{***} ronald.rousseau@kuleuven.be

KU Leuven, Dept. of Mathematics, Celestijnenlaan 200B, B-3001 Leuven (Belgium) & University of Antwerp (UA), IBW, B-2000 Antwerp (Belgium)

ABSTRACT

Networks may be homogeneous in the sense that all nodes are of the same type or they may be heterogeneous in the sense of containing many types of nodes. Once a choice about the types of nodes to be distinguished is made, it is still an open question how to define an acceptable measure of heterogeneity. A proposal to solve this problem is provided in this contribution. In our view the term heterogeneity in a network implies that links between nodes of different type are gauged positively, while links between nodes of the same type should not contribute to a heterogeneity value. Hence a high heterogeneity value refers to a tightly woven net between dissimilar things. We value this property so high that networks without external links receive the same heterogeneity value as homogeneous networks. Concretely, units that determine the value of a heterogeneity measure are links between nodes of different types. These considerations lead to a new measure for heterogeneity derived from a true diversity measure. We claim that we are now able to measure the heterogeneity of networks in a much more precise way than was possible before. An example related to interdisciplinarity is provided. As heterogeneous networks are ubiquitous in the real world, such as in molecular networks, disease networks and trade networks our approach has universal applicability.

INTRODUCTION

Networks may be homogeneous in the sense that all nodes are of the same type or they may contain many types of nodes. A bipartite network linking authors with papers provides an example. Similarly, the nodes in a network of bibliographically coupled or co-cited articles can be attributed to the journal in which they are published.

In this contribution the term network heterogeneity refers to links connecting different types of nodes. Whether two nodes are considered to be of different type depends on the application one has in mind. Once a choice about the types of nodes to be distinguished is made, it is still an open question how to define an acceptable measure of heterogeneity. In this contribution, we propose a solution to this problem for the case of undirected networks. Links between nodes of the same type are called internal links, while links between nodes of different type are called external links. Heterogeneous networks are ubiquitous in the real world; hence a

¹ This work was supported by the National Natural Science Foundation of China (NSFC grant No. 71573225).

precise definition of heterogeneity is required. In this contribution such a measure is proposed.

HETEROGENEOUS NETWORKS AND AN INTERPRETATION AS LAYERED NETWORKS

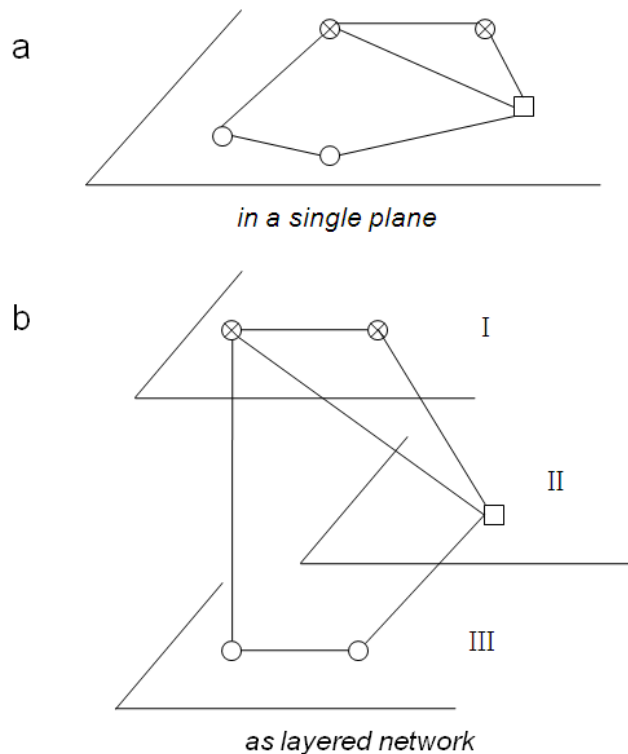
In our view the term heterogeneity in a network implies that external links are gauged positively, while internal links should not contribute. Moreover, we prefer evenly distributed external links above unevenly distributed ones (Hill, 1973). This is a choice we made, admitting that other choices, depending on the application one has in mind can be made. In this contribution a high heterogeneity value refers to a tightly woven net between dissimilar things. Networks without external links receive the same heterogeneity value as homogeneous networks. Indeed, when different types of nodes are never connected the network is actually a disjoint union of homogeneous networks. Concretely, units that will determine the value of a heterogeneity measure are links between nodes of different types. We stress the point that units are not nodes but external links.

Interpretation as a layered network

If there are N types of nodes these may be considered as belonging to N layers, leading to a layered network. Links connecting different types of nodes, i.e. external links, then become links between different layers. In this way the theory of layered networks (Boccaletti et al., 2014) can be applied to heterogeneous networks.

The following example (Fig.1) illustrates how a heterogeneous network can be seen as a layered one.

Figure 1: a Representation in a single plane; b Representation as a layered network



DIVERSITY AS MEASURED THROUGH VARIETY AND BALANCE

We claim that node heterogeneity can be considered as a particular form of diversity, in which links between different types of nodes are the essential constituents. Before proposing a measure for this type of heterogeneity we recall that traditionally diversity is measured through the notions of variety and balance (Magurran, 2003).

The definitions of variety and balance

Variety is the number of non-empty categories to which system elements are assigned. Assuming all other things equal, the greater the variety, the greater the diversity. Balance is a function of the assignment of elements across categories. It answers the question: What is the relative number of items of each type? All else being equal, the more balanced the distribution, the larger the diversity. Variety is a positive natural number as categories are numbered in sequence; balance is a function of fractions summing up to one.

HETEROGENEITY

From now on we use the term heterogeneity when referring to the diversity of links between layers of different types of nodes. Moreover, we follow Hill (1973) and Jost (2006, 2009) requiring that heterogeneity measures should be so-called “true” heterogeneity measures. The main point about these measures is that only when working with true measures it makes sense to discuss heterogeneity in terms of ratios or percentages.

Following Jost’s (Jost, 2006) arguments in favor of true diversity we apply his formula qD for the measurement of heterogeneity of networks choosing $q = 2$. Assuming there are N layers this leads to a Hirschman-Simpson-Herfindahl type of heterogeneity measure, which we denote by HE. If the network itself is denoted by X , we have:

$$HE(X) = \left(\sum_{i=1}^N \sum_{j=i+1}^N p_{ij}^2 \right)^{\frac{1}{1-2}} = \frac{1}{\sum_{i=1}^N \sum_{j=i+1}^N p_{ij}^2} \quad (1)$$

The symbols p_{ij} refer to the relative number of links, among external links, between nodes of type i and of type j . In the case of a homogeneous network X - this means that all nodes are of the same type and hence $N = 1$ - we set $HE(X)$ equal to zero. Similarly, the heterogeneity value of a network without external links is also set equal to zero.

A simple example

We determine the heterogeneity of the example network shown in Fig.1. In this case $N=3$. Relative proportions are: $p_{I,II} = 0.5$, $p_{I,III}=0.25$ and $p_{II,III} = 0.25$. Hence

$$HE(X) = \frac{1}{p_{I,II}^2 + p_{I,III}^2 + p_{II,III}^2} = \frac{1}{\frac{1}{4} + \frac{1}{16} + \frac{1}{16}} = \frac{1}{\frac{6}{16}} = \frac{8}{3}$$

What happens when a peripheral node is added?

By a peripheral node we mean a node which forms a type on its own and is linked to no other or exactly one other node. If a node which forms a type on its own is not linked to any other node, then the heterogeneity measure stays the same.

If a node which forms a type on its own is linked to exactly one other node, then N becomes N+1 and the number of different types of links (M) increases by one, too. We denote the original network by X and the one with one extra node attached by X'. If $p_{ij} = n_{ij}/M$ (where n_{ij} is the number of links between nodes of type i and nodes of type j in network X) then in X' the corresponding p_{ij} becomes $n_{ij}/(M+1) = p_{ij} (M/(M+1))$. Now,

$$HE(X') = \frac{1}{\sum_{i=1}^N \sum_{j=i+1}^N (p'_{ij})^2 + (\frac{1}{M+1})^2} = \frac{1}{\sum_{i=1}^N \sum_{j=i+1}^N (p_{ij} \frac{M}{M+1})^2 + (\frac{1}{M+1})^2} = (\frac{M+1}{M})^2 \frac{1}{\sum_{i=1}^N \sum_{j=i+1}^N (p_{ij})^2 + (\frac{1}{M})^2}$$

$$= (\frac{M+1}{M})^2 \left(\sum_{i=1}^N \sum_{j=i+1}^N (p_{ij})^2 + (\frac{1}{M})^2 \right)^{-1} = (\frac{M+1}{M})^2 \left((HE(X))^{-1} + \frac{1}{M^2} \right)^{-1}$$

EXTENSION TO WEIGHTED LINKS AND TO THE CASE THAT A NODE MAY BE OF DIFFERENT TYPES

When links are valued or weighted positively, formula (1) can still be used, but the p_{ij} become relative weights with respect to the total weight. If a node can be of more than one type it is assumed to belong to several different layers, linked to replica of itself. As space does not permit we omit the technical details for this case.

A REAL-WORLD EXAMPLE

As an illustration we determined the heterogeneity of the reference lists of twelve articles which were studied in an earlier article by Rafols and Meyer (2010) in the context of a study on interdisciplinarity of nanobioscience. We refer to the original publications for bibliographic details of these twelve publications. Here we denote them in the same way as in (Rafols & Meyer, 2010). Table 1 provides a comparison between some indicators and the heterogeneity measure (HE) for these 12 papers. Data for the Stirling index (Stirling, 2007) and the mean linkage strength (S) are taken from (Rafols & Meyer, 2010). Results for the Stirling index are obtained based on the distribution of WoS subject categories of references of references.

Table 1. A comparison between some indicators and our heterogeneity measure

First author of article (year of publication)	#WoS fields	Stirling index (refs of refs)	S: mean linkage strength of bc network	HE
Burgess (2003)	8	0.14	0.050	3.93
Funatsu (1995)	6	0.27	0.054	2.24
Noji (1997)	4	0.15	0.024	5.31
Ishijima (1998)	7	0.18	0.042	7.41
Kikkawa (2001)	8	0.16	0.072	5.45
Kojima (1997)	4	0.24	0.074	4.39
Okada (1999)	4	0.15	0.107	3.13
Sakakibara (1999)	6	0.16	0.029	5.29
Tomishige (2000)	7	0.14	0.104	5.70
Tomishige (2002)	5	0.15	0.113	4.97
Yasuda (1998)	4	0.14	0.039	3.98
Yildiz (2004)	11	0.17	0.065	13.87

Table 2. Pearson correlation between diversity, network coherence, variety of subject categories and heterogeneity

	#WoS fields	Stirling	Coherence	Heterogeneity
# WoS fields	1.00	-0.08	-0.02	0.73
Stirling		1.00	-0.09	-0.17
Coherence			1.00	-0.06
Heterogeneity				1.00

It can rightly be argued that the distribution of references of references is an indicator of interdisciplinarity, yet it is not an indicator of WoS category heterogeneity of references. As such we are not surprised that there does not seem to be a relation between the Stirling index and the HE-measure. It seems though that the number of WoS fields present in the reference list plays a significant role. This is to be expected as in this example ‘heterogeneity’ is characterized by the presence of different WoS fields in the reference list. Yet, even this small example shows that there is no simple one-to-one correspondence between the number of WoS categories present and the HE-value.

DISCUSSION AND CONCLUSION

The practical example we elaborated is related to interdisciplinarity, considered here as a special case of heterogeneity. Yet we stress the point that our approach deals with heterogeneity in general and is not restricted to interdisciplinarity. Heterogeneous networks, composed of different types of objects are ubiquitous in the real world. In other words, our new indicator for heterogeneity has universal applicability.

REFERENCES

- Boccaletti, S., Bianconi, G., Criado, R., et al. (2014). The structure and dynamics of multilayer networks. *Physics Reports*, 544(1), 1-122.
- Hill, M.O. (1973). Diversity and evenness: a unifying notation and its consequences. *Ecology*, 54(2), 427-432.
- Jost, L. (2006). Entropy and diversity. *Oikos*, 113(2), 363-375.
- Jost, L. (2009). Mismeasuring biological diversity: Response to Hoffmann and Hoffmann (2008). *Ecological Economics*, 68(4), 925-928.
- Magurran, A.E. (2003). *Measuring biological diversity*. Wiley-Blackwell.
- Rafols, I., & Meyer, M. (2010). Diversity and network coherence as indicators of interdisciplinarity: case studies in bionanoscience. *Scientometrics*, 82(2), 263-287.
- Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface*, 4(15), 707-719.

CHAPTER 15

Altmetrics



Article-level metrics and the periphery: an exploration of articles by Brazilian authors¹

Iara Vidal Pereira de Souza* and Fabio Castro Gouveia**

*iaravidalps@ufjf.br

Programa de Pós-Graduação em Ciência da Informação, Universidade Federal do Rio de Janeiro, Rua Lauro Muller 455 - 4º andar, Rio de Janeiro, 22290-160 (Brazil)

**fgouveia@fiocruz.br

Museu da Vida, Fundação Oswaldo Cruz, Av. Brasil 4365, Rio de Janeiro, 21045-900 (Brazil)

ABSTRACT

This work-in-progress paper describes an ongoing PhD study that aims to explore article-level metrics from a set of articles published by Brazilian researchers. It is argued that article-level metrics can offer a more nuanced and accurate picture of the influence of a particular work in comparison to journal-level metrics. However, if these new metrics rely on sources that exclude a large part of research from the periphery, they are at risk of simply preserving the present inequalities in the scholarly communication system. In order to understand how article-level metrics are or could be useful to the scientific community in the peripheries, we need to see what metrics are currently available, identify possible biases, and understand their meaning. We aim to contribute to this discussion with a case study focused on exploring a set of both traditional and alternative article-level metrics related to publications authored by Brazilian researchers. So far, few studies analyse article-level metrics for Brazilian publications, and most focus on Brazilian journals instead of researcher's affiliation. Our study will collect articles with DOIs registered by Brazilian researchers at the Lattes Platform, an information system maintained by the national Science, Technologies and Communications ministry. This exploration aims to address the following questions: (a) Which are the main article-level metrics available for journal articles authored by Brazilian researchers? What are the main sources of ALM data for Brazilian publications?; (b) Are there any disciplines, institutions, locations etc. that attract more mentions in the case of Brazilian articles? How do these metrics compare among themselves?; and (c) Do article-level metrics of publications by Brazilian researchers reflect patterns and trends observed in studies with researchers from other countries?

INTRODUCTION

This paper describes an ongoing PhD research that aims to explore article-level metrics from a set of articles published by Brazilian scientists in order to understand their meaning and potential for the evaluation of research done by scientists in (semi)periphery countries.

Article-level metrics (ALM) combine traditional metrics like citations, usage metrics like visualizations and downloads, and so-called "alternative" metrics (or altmetrics) such as mentions on social media, reference managers and news outlets, in order to offer a more nuanced and accurate picture of the influence of a particular work inside and outside academia (Chamberlain, 2013). Recent initiatives such as the San Francisco Declaration on Research Assessment (American Society for Cell Biology, 2012) and the Leiden Manifesto

¹ This work is supported by a grant from the Brazilian agency for graduate education, CAPES.

(Hicks, Wouters, Waltman, de Rijcke, & Rafols, 2015) advocate for the use of ALM for evaluating individual researchers and institutions, substituting journal-level metrics such as the Journal Impact Factor.

For researchers in periphery and semi-periphery countries like Brazil, the use of journal-level metrics in individual evaluation poses specific problems, mainly because the international databases used to calculate these metrics include only a fraction of the titles published in developing countries. In the 80s, Velho pointed out that the Science Citation Index (SCI, the only computerised source for scholarly citations at the time) was “arguably biased in favour of the scientifically central nations” (1986, p. 73). Even though many additions were made to the SCI database (now called Web of Science) in the past 30 years, a large part of the research coming from the peripheries is still excluded from it and from similar sources such as Scopus (Alperin, 2014; Vessuri, Guédon, & Cetto, 2014). If article-level metrics can provide a better picture of the impact of the research published by scientists from the periphery, they could potentially help local researchers, funders and policymakers apply their resources more effectively.

However, there’s no reason to believe article-level metrics are free of the biases already observed in journal metrics. Citations counts are arguably a more accurate representation of the influence of a particular paper than the impact factor of the journal it was published, but if these citations are collected only from international databases that exclude a large part of research from the periphery, they’ll still potentially leave out a large part of the citations actually received by that paper. Similarly, if altmetrics tools are only capable of searching North-American and European sources and/or documents written in English, they might “inadvertently replicate the exclusion of developing countries that has plagued the traditional measures of impact” (Alperin, 2013, p. 20). In short, instead of helping reveal the impact of the research done in the periphery, which is often rendered invisible by current evaluation metrics, a turn to article-level metrics could simply preserve the present inequalities in the scholarly communication system.

We believe that, in order to understand how article-level metrics are or could be useful to the scientific community in the peripheries, we need to see what metrics are currently available, identify possible biases, and understand their meaning. We aim to contribute to this discussion with a case study focused on exploring a set of both traditional and alternative article-level metrics related to publications authored by Brazilian researchers.

Our exploration of these data aims to address the following questions:

- (a) Which are the main article-level metrics available for journal articles authored by Brazilian researchers? What are the main sources of ALM data for Brazilian publications?
- (b) Are there any disciplines, institutions, locations etc. that attract more mentions in the case of Brazilian articles? How do these metrics compare among themselves?
- (c) Do article-level metrics of publications by Brazilian researchers reflect patterns and trends observed in studies with researchers from other countries?

METHODS

In order to obtain a sample of journal articles authored by Brazilian researchers, we’ll collect Digital Object Identifier (DOI) information registered by these researchers at the Lattes Platform (<http://lattes.cnpq.br/>). The Lattes platform is an information system maintained by the national Science, Technologies and Communications ministry where Brazilian researchers

(and researchers from other countries working in Brazil) can record their professional activities, publications, research projects, science communication and outreach efforts, and other information. Having an updated Lattes curriculum vitae is a requirement for many activities in Brazil's scientific community, including but not limited to graduate program admission, job and grant applications, and institutional evaluations. Because of this, data from the Lattes platform offers a reasonably accurate picture of the active scientific community in Brazil. According to a spreadsheet provided at the Platform website (<http://lattes.cnpq.br>, section "Extração de Dados" or Data Extraction), there were a total of 4.559.599 CVs registered in the platform at April 14 2016.

Information provided by researchers to the Lattes Platform is public, and all the curricula are available for consultation on its website (<http://lattes.cnpq.br>). However, since 2015, it is necessary to solve a CAPTCHA in order to view a CV, a measure taken after some websites were found to be using personal data from Lattes for commercial purposes. This poses a challenge for collecting the data necessary for this study, but we are working with tools developed by other Brazilian researchers that facilitate extraction of Lattes data for academic purposes. In order to not overload the Lattes servers, data collection will be preferably done during low usage hours (from 02:00AM to 08:00AM, according to our assessment).

Since 2007, the Lattes platform allows researchers to include articles with DOIs in their curricula, and only CVs that were modified since that date will be searched for DOIs (4,369,958). We have decided to focus on these articles in order to facilitate the process of obtaining metrics (since current ALM tools usually require permanent identifiers), and enable comparisons with previous studies using different countries and/or databases. Besides, by using Lattes as our source we hope to avoid some selection biases that would be present if we chose to gather data from an international database such as Scopus or Web of Science, and to allow for comparisons between papers published in local and international journals. It should be noted, however, that using DOIs might exclude a potentially significant portion of articles by Brazilian researchers, those published in journals that still don't use permanent identifiers. Other studies are looking into ways of collecting altmetrics for articles without DOIs, with some promising preliminary results (Araújo, Murakami, Lara, & Fausto, 2015).

After obtaining these DOIs from the Lattes platform, we will identify the articles they belong to, collecting their titles, authors, journals, discipline, and date of publication. We will also use Lattes to collect information about their authors, such as institutions, location, and education level. Finally, we will gather and analyse metrics related to these articles, such as citations, downloads, and altmetrics data (Twitter, Facebook, Mendeley, news outlets and blog mentions). We are currently in the process of deciding which source(s) to use for collecting ALM data.

ALTMETRIC SOURCES AND BRAZILIAN PUBLICATIONS

There are four main altmetric providers in the market today: ImpactStory, Plum Analytics, Altmetric, and PLOS Article-Level Metrics. ImpactStory offers services for individual researchers, while the others focus on publishers and/or institutions. All of them combine traditional and so-called alternative sources, using different sources and different methods for collecting and aggregating data. Previous investigations on altmetrics using Brazilian publications have used data either from Altmetric (Alperin, 2015; Araújo, 2014; Nascimento & Oddone, 2015) or PLOS ALM (Fausto & Mugnaini, 2014). Others have preferred to gather data directly from Twitter and Facebook (Araújo, Murakami, Lara, & Fausto, 2015).

Alperin (2015) collected article-level data for a sample of articles from SciELO, a Latin American platform for open access journals. Most of the altmetrics were provided by Altmetric.com, with the exception of data from Mendeley and Wikipedia, which were collected via specific scripts. The study found negligible or null coverage levels for SciELO Brazil articles in most of the social media sources analysed, with only Mendeley (18.80% of articles with at least one reader), Twitter (6.03% of articles with at least one Tweet), and Facebook (2.81% of articles with at least one post) presenting coverage levels above 2 per cent. These levels of coverage are generally lower than those reported in previous studies that analyse articles from sources such as Web of Science or Pubmed, among others. For Mendeley, coverage levels are similar to the ones found by previous studies, but SciELO Brazil articles appear to take about two years longer than average to be saved in the reference manager.

Araújo (2014) and Nascimento and Oddone (2015) used Altmetric tools to collect mentions for articles published in different samples of Brazilian Information Science journals. Araújo (2014) found only 6 articles with at least one mention on a sample of 121 articles. Nascimento and Oddone (2015) found 101 articles that had at least one altmetric mention in a sample of articles published by the only five Brazilian Information Science journals covered by the Altmetric Explorer at the time (they do not clarify the total number of articles in their sample). Most of these mentions came from Mendeley (1,001) with Twitter (131) in second and Facebook (25) in third.

Fausto and Mugnaini (2014) used the PLOS ALM Reports tool to analyse article-level metrics related to articles published in PLOS journals by researchers affiliated with Universidade de São Paulo (São Paulo University, USP), the largest university in the country, comparing their performance with that of other articles in PLOS journals from researchers belonging to other Brazilian institutions. They found that researchers located in Brazil authored a little over 2% (1,598) of all the articles published by PLOS journals between 2005 and 2012 (69,306). Their data shows Mendeley (33,733 mentions) and Facebook (14,450 mentions) as the leading sources of altmetric data for their sample. Twitter, which came second in the studies of Alperin (2015) and Nascimento and Oddone (2015), provided 2,284 mentions in this sample, considerably less than Facebook.

A study by Araujo, Murakami, Lara, and Fausto (2015) used a different strategy, focusing on a single Brazilian Information Science journal, Datagramazero, and collecting data directly from Facebook and Twitter via their respective APIs (Application Programming Interface). This particular journal doesn't apply DOIs to its articles, which would make it impossible to use Altmetric tools. Of a total of 441 articles in the sample, 211 articles obtained one or more mentions. The combined number of mentions was 1,164, with an average of 2.63 mentions per article. 84.28% of the mentions came from Facebook, and only 15.72% from Twitter.

The relatively low proportion of Facebook mentions in data provided by Altmetric, when compared to data obtained from PLOS ALM or from the Facebook API is probably due to differences in which data are collected, and how. Altmetric notes in its "Our sources" web page (<https://www.altmetric.com/about-our-data/our-sources/>) that it only counts mentions on public Facebook pages, excluding personal posts. Considering the popularity of Facebook in Brazil and the evidence presented above, it seems that Altmetric could be missing a significant share of mentions to Brazilian journal articles in that social network.

These previous experiences provide useful insights for deciding how to collect article-level data for our study. Based on the above, we have decided to look for alternatives to Altmetric, which would initially be our choice based on its current prevalence in international studies. One possible substitute would be Lagotto, the open source software used by the PLOS Article-Level Metrics tool. Lagotto also serves as the basis for Crossref Event Data (<http://eventdata.crossref.org/>), Crossref's tool for gathering metrics related to any content identified with a DOI, which is currently under development with the official launch planned for the second half of 2016.

We also notice that, so far, most studies of article-level metrics for Brazilian publications have focused on Brazilian journals (Alperin, 2015; Araújo, 2014; Araújo, Murakami, Lara, & Fausto, 2015; Nascimento & Oddone, 2015). We found only one study that used author affiliation to define Brazilian publications (Fausto & Mugnaini, 2014), but it focused on a single institution and it looked only at articles published in PLOS journals. Our study attempts to look at a more representative sample of Brazilian research by using data provided at a national CV platform by researchers themselves, with no institution nor journal restrictions.

NEXT STEPS

At this moment we are finishing our literature review on article-level metrics, which aims to identify issues and questions of interest for our own exploration. For instance, a study of papers published in 2012 and indexed by the Web of Science compared the effect of document characteristics (i.e., discipline, document type, title length, number of pages and references) and collaborative practices on citations and alternative metrics patterns, finding that while both citations and altmetrics increase with the extent of collaboration and the length of the references list, differences arise when looking into other characteristics like discipline – Social Sciences and Humanities papers tend to attract more attention in social media platforms than papers in other fields (Haustein, Costas, & Larivière, 2015). We'll look for these kind of patterns in our own data in order to understand how the use of Brazilian articles might reflect them.

In the next few months we will start the collection and analysis of DOI data from the Lattes platform. This analysis should provide information, such as which journals and/or areas are more prevalent in our sample, that will help us decide which source(s) of ALM data, especially those related to citations and usage, would be more useful to our research. Then, a pilot study with a smaller sample will serve to identify possible issues with our strategy before working in a greater scale.

REFERENCES

- Alperin, J. P. (2013). Ask not what altmetrics can do for you, but what altmetrics can do for developing countries. *Bulletin of the American Society for Information Science and Technology*, 39(4), 18–21. Retrieved from <http://doi.wiley.com/10.1002/bult.2013.1720390407>
- Alperin, J. P. (2014). South America: Citation databases omit local journals. *Nature*, 511(7508), 155–155. Retrieved March 27, 2016, from <http://www.nature.com/doifinder/10.1038/511155c>
- Alperin, J. P. (2015). Geographic variation in social media metrics: an analysis of Latin American journal articles. *Aslib Journal of Information Management*, 67(3), 289–304.

- Retrieved August 4, 2015, from <http://www.emeraldinsight.com/doi/abs/10.1108/AJIM-12-2014-0176>
- American Society for Cell Biology. (2012). The San Francisco Declaration on Research Assessment: putting science into the assessment of research. ASCB. Retrieved from <http://am.ascb.org/dora/>
- Araújo, R. F. de. (2014). Cientometria 2.0, visibilidade e citação: uma incursão altmétrica em artigos de periódicos da ciência da informação. *ENCONTRO BRASILEIRO DE BIBLIOMETRIA E CIENTOMETRIA, 4., 2014, Recife* (p. 8). Recife: UFPE. Retrieved from <http://www.brapci.inf.br/article.php?dd0=0000014387&dd90=dd87dc69fb>
- Araújo, R. F. de, Murakami, T. R. M., Lara, J. L. de, & Fausto, S. (2015). Does the Global South have altmetrics? Analyzing a Brazilian LIS journal. Presented at the 15th International Conference on Scientometrics and Informetrics, Istanbul. Retrieved July 4, 2015, from https://www.researchgate.net/publication/279443581_Does_the_Global_South_have_altmetrics_Analyzing_a_Brazilian_LIS_journal
- Chamberlain, S. (2013). Consuming Article-Level Metrics: Observations and Lessons. *Information Standards Quarterly, 25*(2), 4. Retrieved from <http://www.niso.org/publications/isq/2013/v25no2/chamberlain/>
- Fausto, S., & Mugnaini, R. (2014). Beyond traditional metrics at the University of São Paulo: scientific production in the PLOS journals. *Context Counts: Pathways to Master Big and Little Data*. Presented at the International Conference on Science and Technology Indicators, 19th, Leiden: Leiden University. Retrieved March 20, 2015, from <http://pt.slideshare.net/sfausto/sti2014-posteruspplosfinal>
- Haustein, S., Costas, R., & Larivière, V. (2015). Characterizing Social Media Metrics of Scholarly Papers: The Effect of Document Properties and Collaboration Patterns. *PLoS ONE, 10*(3), e0120495. Retrieved March 18, 2015, from <http://dx.doi.org/10.1371/journal.pone.0120495>
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature, 520*(7548), 429–431. Retrieved June 17, 2015, from <http://www.nature.com/doi/10.1038/520429a>
- Nascimento, A. G. do, & Oddone, N. E. (2015). Uso de altmetrics para avaliação de periódicos científicos brasileiros em Ciência da Informação. *Ciência da Informação em Revista, 2*(1). Retrieved July 4, 2015, from <http://www.seer.ufal.br/index.php/cir/article/view/1745>
- Velho, L. (1986). The „meaning“ of citation in the context of a scientifically peripheral country. *Scientometrics, 9*(1–2), 71–89. Retrieved from <http://link.springer.com/10.1007/BF02016609>
- Vessuri, H., Guédon, J.-C., & Cetto, A. M. (2014). Excellence or quality? Impact of the current competition regime on science and scientific publishing in Latin America and its implications for development. *Current Sociology, 62*(5), 647–665. Retrieved November 30, 2014, from <http://csi.sagepub.com/content/62/5/647>
-



Can we use altmetrics at the institutional level? A case study analysing the coverage by research areas of four Spanish universities¹

Daniel Torres-Salinas^{*}, Nicolas Robinson-Garcia^{**} and Evaristo Jiménez-Contreras^{***}

^{*}torressalinas@gmail.com

Universidad de Navarra and Universidad de Granada (EC3metrics and Medialab UGR), Granada (Spain)

^{**}elrobin@ingenio.upv.es

INGENIO (CSIC-UPV), Universitat Politècnica de València, Valencia (Spain)

^{***}evaristo@ugr.es

Universidad de Granada, Dpto.. Información y Comunicación (Grupo EC3 and EC3metrics), Granada (Spain)

INTRODUCTION

Social media based indicators or *altmetrics* have been under scrutiny for the last seven years. Their promise as alternative metrics for measuring scholarly impact (Priem, Piwowar, & Hemminger, 2012) is still far from becoming a reality (Torres-Salinas, Cabezas-Clavijo, & Jiménez-Contreras, 2013). Issues with regard to their meaning (Sud & Thelwall, 2014), potential use as alternative or complements to citation indicators (Bornmann, 2014; Costas, Zahedi, & Wouters, 2015a), data collection inconsistencies (Robinson-García, Torres-Salinas, Zahedi, & Costas, 2014) or diversity of sources (Thelwall, Haustein, Larivière, & Sugimoto, 2013) are still under development. In this regard, three lines of work can be observed in regard to altmetric studies: 1) the relation between citations and altmetric indicators (Costas et al., 2015a), 2) their meaning as impact indicators (i.e., Bornmann, 2014; Haustein, Bowman, & Costas, 2015), and, 3) coverage and diversity of social media sources (i.e., Haustein, 2016).

Altmetric indicators have been intimately related since their conception with commercial interests (Bornmann, 2014) and many data providers offering social media metrics have been developed. Here Altmetric.com has positioned itself as the most spread and used data sources for altmetric studies. Altmetric.com has the advantage of providing detailed data from a variety of social media platforms with regard to mentions, readership, etc. of scientific papers. It also offers an aggregated indicator or ‘altmetric’ score which is based on a weighted sum of values based on the presence of a given article in different social media. Contrarily to its competitors, it has the advantage of maintaining the ‘history’ of each altmetric indicator to an article, thus ending with a classical limitation of this type of indicators: their volatility (Costas et al., 2015a).

Up to now, most studies have focused on the understanding of the nature and relation of altmetric indicators with citation data. Few papers have analysed research profiles based on altmetric data. Most of these have related to researcher profiles and the expansion of these tools among researchers. For instance, (Haustein et al., 2014) surveyed participants of the STI2012 Conference to learn the spread on the use of these tools by researchers. With a

¹ Nicolas Robinson-Garcia is currently supported by a Juan de la Cierva-Formación Fellowship from the Spanish Ministry of Economy and Competitiveness.

similar aim but different approach, (Torres-Salinas & Milanés-Guisado, 2014) analysed the presence in social media of the most prolific authors in the *Profesional de la Información* journal. Finally, (Martín-Martín, Orduna-Malea, & Delgado López-Cózar, 2016) created a portal including altmetric and citation scores for profiling researchers according to the source from where data is retrieved. However, no study has been found profiling larger units of research such as institutions, especially universities. Despite how common it is to find traditional bibliometric studies analysing institutional research performance, currently there is no study focused on the utility, consistency and, especially the thematic coverage of altmetrics at the institutional level.

This paper aims at exploring the coverage of the Altmetric.com database and its potential use in order to show universities' research profiles in relationship with other databases. Specifically, our objectives are the following:

1. Analyse the coverage of Altmetric.com at the institutional level and verify its validity as a data source for obtaining alternative metrics derived from the research activity of universities in comparison with those from the Web of Science. For this, we will work with a small sample of four Spanish universities with different characteristics.
2. Analyse coverage differences when obtaining bibliometric profiles from Altmetric.com and Web of Science. In some studies a higher coverage of the Social Sciences and Humanities has been reported, suggesting the potential of altmetric indicators in these areas (Costas, Zahedi, & Wouters, 2015b).

MATERIAL AND METHODS

In this paper we will analyse a sample of four different Spanish universities. These universities present different sizes and research publication specialization as well as being located in different regions of the country. Table 1 shows a description of the characteristics of each university, the main field in which they are specialised, foundation year, and total number of publications produced in 2014 according to the Web of Science database.

Table 1. Description of the four universities analysed in this study

<i>University</i>	<i>Size Staff Full Time</i>	<i>Type and main field</i>	<i>Foundation Year</i>	<i>Region & City</i>	<i>Nº Web of Science Docs 2014*</i>
University of Granada	2399	Multidisciplinary	1531	Andalucía Granada	2387
University Pompeu Fabra	288	Specialized Medicine/Biology	1990	Cataluña Barcelona	1029
Polytechnic University Valencia	1847	Specialized Engineering	1971	Comunidad Valenciana Valencia	1753
University Carlos III	539	Specialized Social Sciences	1989	Comunidad Madrid Madrid	810

All publications for the 2014 year indexed in the Web of Science citation indexes (SCI, SSCI and H&ACI) as articles, reviews, notes or letters for our sample of institutions were retrieved in November 2015. A total of 5922 records was retrieved. Records from the Web of Science

provide among other information the DOI number of each publication, although this is not necessarily provided for all documents. The DOI number is very important, as it will allow us to query the Altmetric.com API for each document (Robinson-García et al., 2014). A total of 5547 papers from our data set included DOI (93% of the total share). All universities showed a similar share of documents with DOI number except for the University of Granada, where the percentage went down to 91%. Finally, we retrieved all altmetric data available at the moment for our set of papers with DOI.

Publications from the Web of Science have been assigned to four broad research areas: "Engineering & Technology", "Humanities & Arts", "Science" and "Social Sciences". These areas have been constructed by aggregating Web of Science Subject Categories. Then, the share of records with a score assigned by Altmetric.com (as defined above) has been computed. The range of such score for our data set was of 928.

The data set employed for this study is publicly available for reproducibility purposes at <https://dx.doi.org/10.6084/m9.figshare.3120946.v3>. This data set includes all records retrieved from the Web of Science as well as social media metrics and score retrieved from Altmetric.com along with their assigned research area.

RESULTS

Table 1 shows the distribution of publications by research area in order to learn the research profile of each university. The University of Granada is the largest university of the four analysed with up to 2387 papers published in 2014, followed by Polytechnic University of Valencia, Pompeu Fabra University and Carlos III University. Regarding their research profile, we first must note that Science in all cases is the area with the highest share of output. The only exception is Carlos III University, where Engineering & Technology represent 51% of the total share followed by 45% of Science. Engineering has a large presence also in the Polytechnic University of Valencia. Social Sciences reach their highest presence in Carlos III University (20%), followed by Pompeu Fabra University (15%). Humanities & Arts is the area with the least output in the Web of Science, never reaching 10% of the total share of each university.

Altmetric.com covered 5922 records from the Web of Science, representing 36% of our data set (that is, 5922 publications where found to have mentions in social media and had a calculated 'altmetric' score). This share varies considerably depending on the university and research area under consideration. Figure 1 shows the share of Web of Science documents in our data set which have received at least a mention in any of the social media metrics retrieved from Altmetric.com by university. As observed, Pompeu Fabra University as the university best covered by altmetric data: 67% of all its publications received mentions in social media. The other three universities have well under 50% of their total output mentioned in social media, with values between 30-40% for Granada and Valencia, and 23% for Carlos III.

Table 2. Publication profile of four Spanish universities according to the number of papers indexed in the Web of Science in 2014

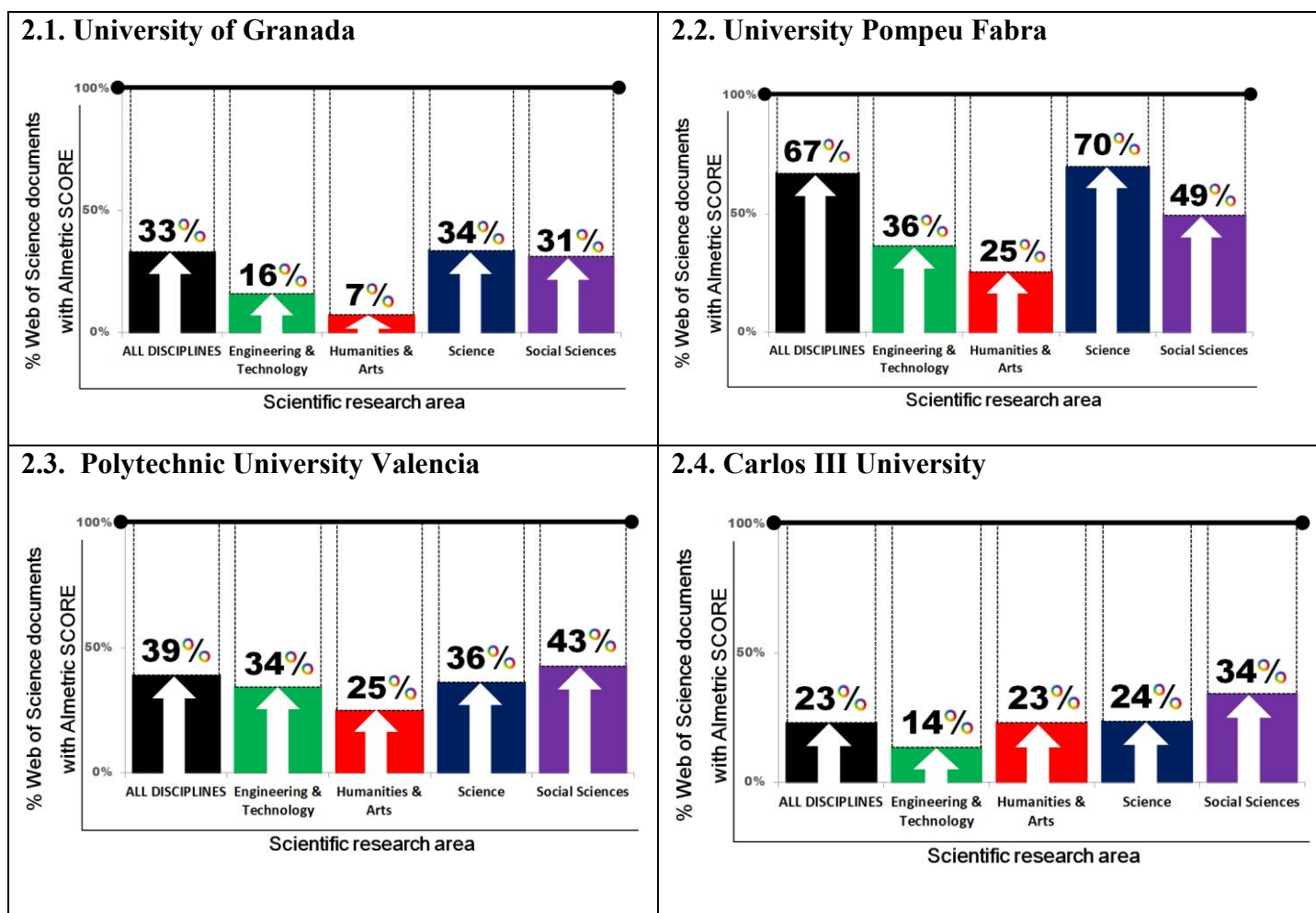
Research Area	University of Granada		Pompeu Fabra University	
	N° WoS Documents	% WoS Documents	N° WoS Documents	% WoS Documents
Engineering & Technology	483	20%	107	10%
Humanities & Arts	84	4%	59	6%
Science	1896	79%	815	79%
Social Sciences	290	12%	156	15%
Total (without duplicates)	2387	100%	1029	100%

Research Area	Polytechnic University Valencia		Carlos III University	
	N° WoS Documents	% WoS Documents	N° WoS Documents	% WoS Documents
Engineering & Technology	779	44%	413	51%
Humanities & Arts	16	1%	39	5%
Science	1272	73%	365	45%
Social Sciences	103	6%	164	20%
Total (without duplicates)	1753	100%	810	100%

If we consider the altmetric coverage by research area and university (figure 1, Science and Social Sciences have a similar coverage of altmetric mentions in the case of the University of Granada (34% and 31% respectively). The area of Science is the best covered for Pompeu Fabra (up to 70%), the only area where the share of mentioned papers surpasses 50%. It is followed by Social Sciences (49%). In the case of the Polytechnic University of Valencia and Carlos III University, Social Sciences is the area best covered. Humanities & Arts is again the research area less well-covered with the exception of Carlos III, where Engineering & Technology show the lowest values.

Figure 2 introduces the altmetric score as a proxy of the intensity of mentions in social media to publications. Here the prevalence of Science in Pompeu Fabra and Granada is evident. Again a pattern can be observed with regard to the areas which receive more mentions in social media: Science first, Social Sciences second, then Engineering & Technology, and Humanities & Arts. In this case, the exception can be found in Polytechnic University of Valencia, where the second area with the highest intensity of social media metrics is Engineering & Technology and not Social Sciences.

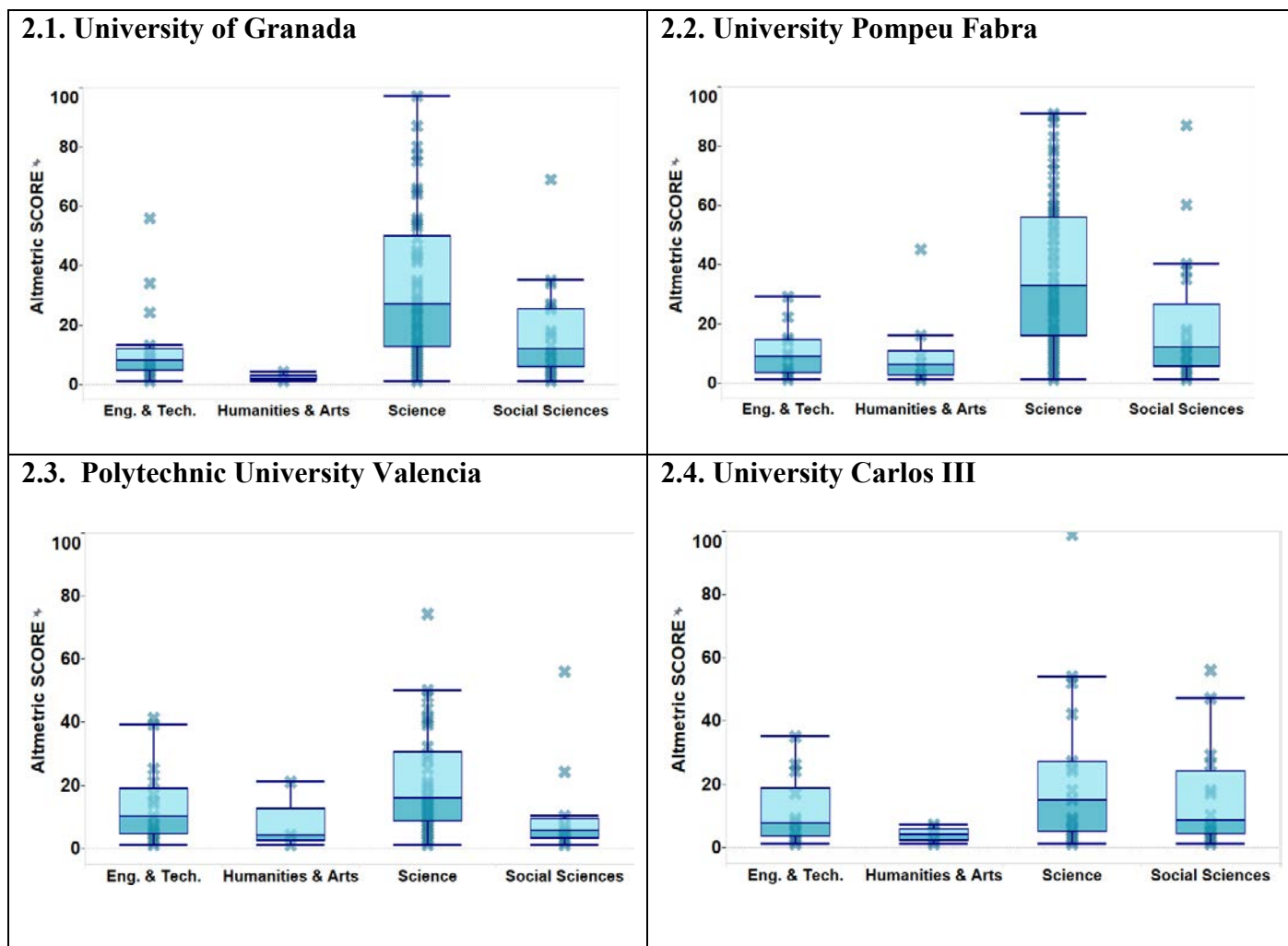
Figure 1. Coverage by fields of altmetric indicators by areas based on the number of documents published by four Spanish universities according to the Web of Science in 2014



DISCUSSION AND CONCLUDING REMARKS

This paper presents a first attempt at analysing altmetric indicators at the institutional level. Specifically, our goal was to analyze to what extent did Altmetric.com showed an adequate and homogeneous coverage among research areas. First, we observe a low coverage of altmetric indicators with only 36% of all documents retrieved from the Web of Science having an ‘altmetric’ score. We only find reasonable levels of coverage for Pompeu Fabra University, where 67% of all documents had altmetric mentions. This university represents a different profile to that of the other three universities analysed, confirming that this university represents an outlier of the Spanish University system (Robinson-García, Rodríguez-Sánchez, García, Torres-Salinas, & Fdez-Valdivia, 2013).

Figure 2. Boxplot with the altmetric score distribution for four Spanish universities by areas in 2014



One of the main factors that may contribute to such low coverage could be the use of the DOI number as reference for querying the Altmetric.com API. One of the first problems encountered here and in any other altmetric study (Robinson-García et al., 2014), is the reliance on DOI numbers to retrieve social media mentions, this assumes a necessary loss of information. First, we must note that not all papers indexed in the Web of Science include DOI (7% of the records in our data set did not include a DOI number). Second and more importantly, we must stress that Altmetric.com uses the DOI for its searches but that 1) not all social media mentions directed to a publication include the DOI number (for instance, Twitter links could use not normalized web links), and 2) not all mentions are directed to the journal article but to other versions of the same publication (i.e., post-refereed versions uploaded to a repository). This translates in a low coverage of altmetric indicators which may be misleading.

Second, we observe that for the four universities analysed, the area of Science shows higher 'altmetric' scores than the rest of the research areas. Science is also the area best covered for three of the four universities. Only in the case of the Carlos III University, Social Sciences are

best covered. Despite this fact, we do not observe for any of the universities analysed a predominance of altmetric data for the areas of Social and Humanities & Arts as suggested elsewhere (Costas et al., 2015a). We could speculate that such differences between our results and previous studies could be due to the fact that none of the studied universities belongs to an English-speaking country.

However, further research is needed to gain more insight as to what is occurring. In this sense, we propose the following lines of work:

1. *The national factor of research areas.* It is necessary to verify if there is a national or linguistic factor influencing the coverage of altmetric indicators, especially in the areas of Social Sciences and Humanities & Arts. Another explanation could be that these areas are represented in social media by small communities, limiting their capacity to disseminate their research papers in social media.

2. *The use of social media by researchers.* It is important to note the influence that the dissemination of research papers by the authors themselves may play in this process. It would be interesting to analyse if universities with an active academic staff in social media could lead to a better coverage of altmetric indicators. Also the contrary should be considered. Are universities with personnel with no presence in social media worse covered by altmetric data?

3. *Collaboration networks.* Another factor that may affect coverage by areas could be international collaboration rates. Areas with more authors per paper could be better covered by social media.

Finally, considering the low coverage of altmetric data at the institutional level, it could be interesting for research policy makers to consider the development of guidelines and best practices guides to ensure that researchers disseminate adequately their research findings through social media, emphasizing the use of normalised identifiers (DOI numbers, ArXiv, PubMedID) in order to ensure the recollection of such metrics.

ACKNOWLEDGMENTS

The authors would like to thank Matthew Spehlmann and Altmetric.com for the collection of the data.

REFERENCES

- Bornmann, L. (2014). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics*, 8(4), 895–903.
- Costas, R., Zahedi, Z., & Wouters, P. (2015a). Do ‘altmetrics’ correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology*, 66(10), 2003–2019.
- Costas, R., Zahedi, Z., & Wouters, P. (2015b). The thematic orientation of publications mentioned on social media: Large-scale disciplinary comparison of social media metrics with citations. *Aslib Journal of Information Management*, 67(3), 260–288.

- Haustein, S. (2016). Grand challenges in altmetrics: heterogeneity, data quality and dependencies. *Scientometrics*, 1–11.
- Haustein, S., Bowman, T. D., & Costas, R. (2015). Interpreting ‘altmetrics’: viewing acts on social media through the lens of citation and social theories. Retrieved from <http://arxiv.org/abs/1502.05701>
- Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (2014). Coverage and adoption of altmetrics sources in the bibliometric community. *Scientometrics*, 101(2), 1–19.
- Martín-Martín, A., Orduna-Malea, E., & Delgado López-Cózar, E. (2016). *The Role of Ego in Academic Profile Services: Comparing Google Scholar, Researchgate, Mendeley, and Researcherid*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2745892
- Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. Retrieved from <http://arxiv.org/abs/1203.4745>
- Robinson-García, N., Rodríguez-Sánchez, R., García, J. A., Torres-Salinas, D., & Fdez-Valdivia, J. (2013). Análisis de redes de las universidades españolas de acuerdo a su perfil de publicación en revistas por áreas científicas. *Revista Española de Documentación Científica*, 36(4), e027.
- Robinson-García, N., Torres-Salinas, D., Zahedi, Z., & Costas, R. (2014). New data, new possibilities: exploring the insides of Altmetric. com. *El Profesional de La Información*, 23(4), 359–366.
- Sud, P., & Thelwall, M. (2014). Evaluating altmetrics. *Scientometrics*, 98(2), 1131–1143.
- Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do Altmetrics Work? Twitter and Ten Other Social Web Services. *PLoS ONE*, 8(5), e64841.
- Torres-Salinas, D., Cabezas-Clavijo, Á., & Jiménez-Contreras, E. (2013). Altmetrics: New Indicators for Scientific Communication in Web 2.0. *Comunicar*, 21(41), 53.
- Torres-Salinas, D., & Milanés-Guisado, Y. (2014). Presencia en redes sociales y altmétricas de los principales autores de la revista. *El Profesional de la Información*, 23(4), 367–372.



Enriching the knowledge of altmetrics studies by exploring social media metrics for Economic and Business Studies journals

Kaltrina Nuredini^{*}, Isabella Peters^{*(*)}

^{*}k.nuredini@zbw.eu; i.peters@zbw.eu

Web Science, ZBW - Leibniz Information Centre for Economics, Düsternbrooker Weg 120, Kiel, 24105

^(*)Kiel University, Christian-Albrechts-Platz 4, 24118 Kiel
(Germany)

ABSTRACT

We present a case study of articles published in 30 journals from Economics and Business Studies (EBS) by using social media metrics from Altmetric.com. Our results confirm that altmetric information is significantly better present for recent articles. The Top 3 most used altmetric sources in EBS-journals are Mendeley, Twitter, and News. Low but positive correlations ($r=0.2991$) are identified between citation counts and Altmetric Scores on article level but they increase on journal level ($r=0.614$). However, articles from highly cited journals do neither receive high online attention nor are they better represented on social media.

INTRODUCTION

Nowadays, scholarly journal evaluation and selection by citation indexes (Garfield, 1972) appears to be a debatable story because of the vast amount of studies confirming problematic implications of citation-based indicators (Seglen, 1997), and the rising resistance against inadequate use of the journal impact factor (see e.g., the Leiden Manifesto¹). In response, the introduction of social media tools has led to various social media metrics which are used as indicators for research assessments (Priem et al., 2010; Wouters & Costas, 2012). Altmetrics use sources from the Social Web such as Twitter, Facebook, or reference management tools to quantify the impact of scholarly publications on social media users and can appear more rapidly than citations. Haustein (2016) emphasizes that any metric whether is citation or social media based, has to be wisely chosen depending on the assessment aim. Thus, altmetrics and citation counts are two different measures (Costas et al., 2015). Peters et al. (2014) observed that by using only a subset of social media metrics for research evaluation the results might not correctly reflect the impact of the publications on users. Many studies such as Thelwall et al. (2013) and Costas et al. (2015) reveal that altmetric indicators are associated with citation counts in several disciplines (e.g., biomedical and health sciences, social sciences and humanities, mathematics, life and earth sciences). Altmetrics found application in various fields, e.g., in showcasing scholarly works (i.e. ImpactStory.com).

Also, libraries became interested in using altmetrics data to facilitate filtering of publications and providing context information to publications. It is reasonable for libraries to know which aspects can be implemented in a reasonable way, where sufficient data is available for valid analyses, what altmetrics window (analogous to the citation time window) should be used, and which altmetrics aggregator is the best choice for the goals set.

¹ <http://www.leidenmanifesto.org>

Therefore, contributions to the enhancement of the methodology used for studies and for implementations of social media metrics in real-world applications are needed in order to avoid misinterpretation of indicators in a specific discipline and creation of unintended peripheries by un-reflected use of alternative, as well as traditional, scholarly metrics. This case study in Economics and Business Studies (EBS) literature will enhance the discussion of altmetrics and can act as starting point for studies in similar environments. We complement the knowledge on disciplinary peculiarities of altmetrics in order to enhance methodology and interpretation of altmetrics studies in the future.

By using social media metrics provided by the aggregator Altmetric.com we show, for example, how to limit temporal biases in sample creation and what questions to ask when results are set. In the long run, such case studies may help researchers to effectively disseminate or easily evaluate scientific publications since they know which tools are mostly used in what disciplines by what people for which purposes.

The study aims at answering following specific research questions:

RQ1: What is the coverage of journals from EBS in Altmetric.com?

RQ2: Which are the most used altmetric sources for publications from EBS and therefore work best for providing altmetric indicators?

RQ3: Do altmetrics indicators relate to citation counts of publications?

RQ4: Is there any relation between impact factors and the score numbers aggregated by Altmetric.com to reflect importance of journals?

METHODS AND DATA²

The case study on EBS relies on altmetric data provided by the social media metrics aggregator Altmetric.com and on citation data provided by Thomson Reuters' Web of Science (WoS). Altmetric.com collects information for research output found online from specified sources such as social media platforms, traditional media, and online reference managers. In contrast to WoS Altmetric.com looks for both, sources that are related to scholarly content as well as references that rather stem from mainstream media (like popular news outlets). However, analyzed sources need to have APIs to be included in Altmetric.com's index (Robinson-García et al., 2014).

Several studies showed before that the social reference manager Mendeley is a vital reservoir for altmetric data that correlates moderately well with citation counts (amongst others: Zahedi et al., 2015; Thelwall et al., 2013). When working with Altmetric.com it has to be kept in mind, however, that although Altmetric.com retrieves and displays Mendeley reader counts for each available DOI, only those DOIs are saved in the Altmetric.com for which at least one other social media metric (such as Twitter, news, etc.) has been found. Mendeley is not included in the Altmetric Score of Altmetric.com³. Hence, some studies working with data from this provider exclude Mendeley from their analyses (e.g., Costas et al., 2015).

²We thank Fran Davies from Altmetric.com for providing altmetric data and Stefanie Haustein and Vincent Larivière from the Université de Montréal & the Observatoire des Sciences et des Technologies (UQAM) for citation data.

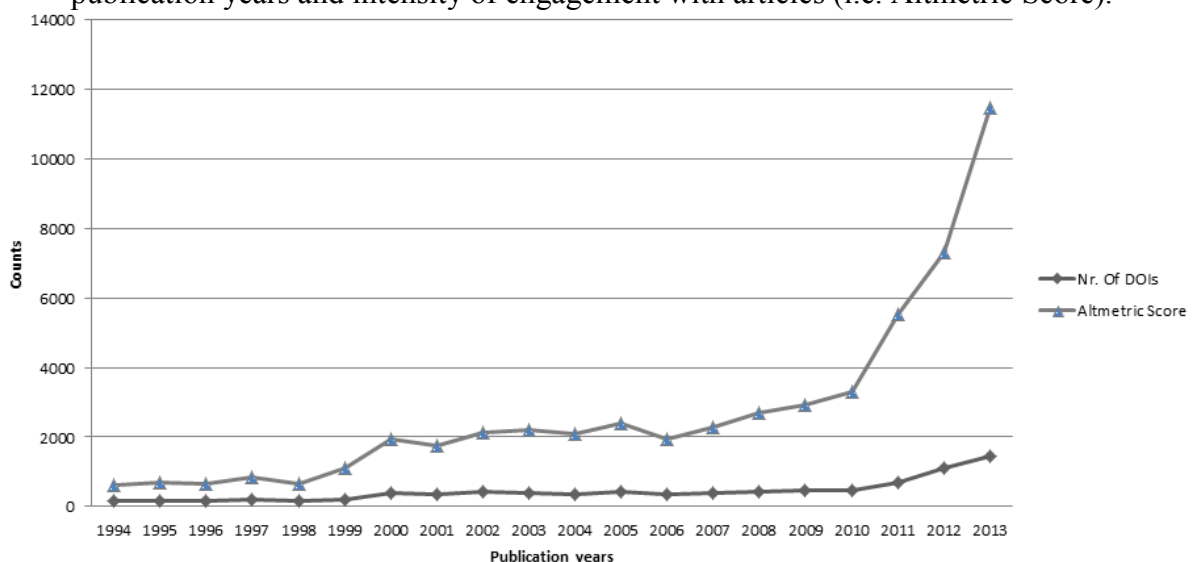
³<https://help.altmetric.com/support/solutions/articles/6000060969-how-is-the-altmetric-attention-score-calculated->

Dataset 1: Coverage of Publications from EBS on Social Media Platforms

We created Dataset I in order to compare altmetrics with citation data of EBS articles and to find the best time span for journal publications with a sufficient amount of altmetrics data. This dataset contains articles that are published in the Top 30 EBS-journals (see Nuredini & Peters, 2015). The data selection is built on journals from the Handelsblatt⁴ journal ranking of which 15 come from Business Studies (Handelsblatt in 2012⁵) and the other half is from Economics (from the Handelsblatt 2010⁶). The dataset is composed of 51,473 DOIs and the articles are restricted to the publication years 1994-2013.

Social media metrics from Altmetric.com were requested for articles of those 30 journals and the search was conducted via journal names. On 27.11.2015 the data came on bulk with a total set of 13,597 DOIs. To filter the year of the publications (1994-2013) the DOIs from Altmetric.com have been matched with the 51,473 DOIs from Nuredini and Peters (2015) resulting in 8,763 DOIs forming Dataset 1 (see Table 1).

Figure 1. Dataset 1: Coverage of DOIs (n=8,763) on social media platforms across 20 publication years and intensity of engagement with articles (i.e. Altmetric Score).



Altmetrics data are present for a bigger share of articles published in recent years (Figure 1). From the publication year 2011 onwards every year more than 10% of the DOIs searched obtained altmetric attention, so it can be concluded that from 2011 there is a considerable and steadily increasing amount of EBS literature available on social media platforms. This temporal bias in altmetric indicators has been already mentioned in earlier studies (e.g., Costas et al., 2015). Moreover, the engagement rates per publication have significantly increased since 2011. Hence, to make adequate use of social media metrics in the field of EBS only publications published from the year 2011 onwards should be considered for further analyses.

⁴ handelsblatt.com

⁵ https://docs.google.com/spreadsheets/d/1GaU_tSl3kC2FtE7xYnSEiktSX9DUKei_qhQKipqyJQ4/pub?output=html

⁶ <http://tool.handelsblatt.com/tabelle/?id=33>

Dataset 2: Comparison of Altmetric Indicators with Citation Counts and Journal Metrics

Given that the analysis of Dataset 1 revealed a strong bias in altmetrics counts towards recent publications a second dataset was compiled. In order to reduce temporal biases in the comparison of citation numbers from WoS (2014 citation counts), i.e. citation delay bias, and altmetrics counts, i.e. social media uptake bias, the study will only analyze articles published from 2012 until 2014.

The Crossref API was queried by ISSN for retrieving the article DOIs of Dataset 1 but restricted to the publication years 2012-2014. This resulted in a total number of 9,045 articles. Then, the list of DOIs from Crossref was matched with the Altmetric.com data and the citation data obtained from WoS. The matching resulted in 3,466 DOIs having social media metrics and 7,410 DOIs found in WoS of which 6,966 have at least one citation (see Table 1).

Table 1. Quantitative description of datasets 1 and 2.

Dataset	Publication Years	Number of DOIs via Crossref	Number of DOIs found in Altmetric.com	Number of DOIs found in WoS with >0 citations
Dataset 1	1994 - 2013	51,473	8,763*	-
Dataset 2	2012 - 2014	9,045	3,466**(*)	6,966
			Number of DOIs having Altmetric and Citation counts	
			3,275***	

*used in RQ1, **used in RQ2, ***used in RQ3, (*) used in RQ4

RESULTS

RQ1: Journal coverage in Altmetric.com

In Table 2, the coverage of articles in Altmetric.com is shown per journal. Overall, 38% of 9,045 searched DOIs are covered, meaning that those 3,466 publications have been mentioned at least once on social media platforms. The highest share of articles represented in the database of Altmetric.com is found for the Quarterly Journal of Economics, where more than two thirds of the published articles have social media metrics. The Journal of Business Research has the highest number of DOIs available but engagement with its articles in the online world is rather low (only 14% of DOIs have altmetrics). The American Economic Review (AER) has the second highest number of articles published and also the highest number of DOIs found in Altmetric.com. But in terms of coverage AER is only on rank 12 (with only 38% of all articles found in Altmetric.com). This is a remarkable result since both journals allow online access, and what is more, AER even has social media buttons integrated into its web pages; functionalities expected to drive users sharing articles within their social media accounts.

When comparing the availability of social media metrics per year the analysis revealed that coverage is steadily increasing for recent publication years: 33% of the DOIs published in 2012 were found in Altmetric.com, 41% from 2013 and 42% from 2014.

Table 2 also shows the journals' Impact Factors (IF) from the 2014 edition of Journal Citation Reports (JCR). Moreover, for each journal the Altmetric Score is displayed, which sums up all social media metrics for each article. The Altmetric Score is defined from Altmetric.com by quantity (the higher the attention, the higher the score) and quality (different social media sources differently impact the score)⁷.

Table 2. Journal metrics ranked according to coverage of articles found in Altmetric.com.

Journals	# DOIs found in Crossref	# DOIs found in Altmetric.com	Coverage of articles in Altmetric.com	Impact Factor (IF)	Altmetric Score
Quarterly Journal of Economics	221	147	67%	6.654	3135
Journal of Health Economics	392	219	56%	2.579	3021
Journal of Consumer Research	411	214	52%	3.125	7985
Economic Journal	371	193	52%	2.336	3561
American Political Science Review	257	124	48%	3.688	2158
Review of Economic Studies	197	94	48%	4.038	102
Journal of Marketing	202	92	46%	3.938	1162
International Organization	180	79	44%	3.019	108
Journal of Finance	446	193	43%	5.424	1287
Administrative Science Quarterly	195	84	43%	3.333	171
Journal of Political Economy	185	78	42%	3.593	626
American Economic Review	1087	415	38%	3.673	5255
Journal of Labor Economics	143	53	37%	1.893	476
Journal of Econometrics	698	238	34%	1.600	500
Econometrica	388	126	32%	3.889	606
Management Science	852	225	26%	2.482	1978
Journal of Marketing Research	292	68	23%	2.256	1721
Academy of Management Journal	375	80	21%	6.448	1502
Journal of Financial Economics	553	117	21%	4.047	855
Journal of Monetary Economics	361	74	20%	1.726	374
Information Systems Research	260	52	20%	2.436	245
The Annals of Statistics	420	66	16%	2.180	173
European Economic Review	496	74	15%	1.444	557
Journal of Business Research	1495	207	14%	1.480	904
Journal of Accounting and Economics	204	28	14%	2.724	95
Academy of Management Review	216	23	11%	7.475	228
International Economic Review	224	22	10%	1.210	319
Games and Economic Behavior	612	49	8%	1.067	101
Journal of Economic Theory	553	43	8%	1.033	191
Journal of Business and Economic Statistics	282	16	6%	2.241	49

⁷ <https://help.altmetric.com/support/solutions/articles/6000059309-about-altmetric-and-the-altmetric-score>

RQ2: Best providers of altmetric sources

Table 3 displays 14 sources for which altmetrics for the DOIs of Dataset 2 have been collected by Altmetric.com. By summing up the usage numbers that every social media source tracked by Altmetric.com has achieved the “Total count of Altmetric events” is calculated. For example, we summed up the number of tweets of Twitter from every article in our dataset resulting in 21,716 counts in total.

Table 3. Sources found in Altmetric.com for 3,466 DOIs.

Altmetric Source	Total Count of Altmetric Events	Number of DOIs found for this altmetric source	Mean Events per Publication
Mendeley	159,354	3,333	47.81
Twitter	21,716	3,080	4.95
CiteUlike	329	258	1.27
Blogs	833	577	1.44
Wikipedia	126	102	1.23
News	1,186	421	2.81
Policy Documents	183	165	1.10
Facebook	581	398	1.46
Google+	198	122	1.62
Weibo	131	86	1.52
Reddit	71	42	1.69
F 1000	4	4	1
Peer review	4	4	1
Pinterest	7	7	1

As observed before, 77% of articles from our dataset have altmetric readership counts from Mendeley; hence it is the source providing most altmetric counts for EBS publications. Twitter has 88% of the found DOIs, News sum up to 34%, mentions in Blogs to 24%, Facebook shares are 16%, and other sources are below 15% each. Interestingly, although most of the DOIs from the dataset have been found on Mendeley there is still a small share of 3.84% of DOIs which could not be found via this social reference tool but via other services.

Nevertheless, besides that Mendeley accumulates more metrics Twitter, Blogs, Facebook, and news are identified as sources of substantial altmetric data in EBS. The coverage of Top 30 journals in 13 social media sources is shown in Figure 2. We have excluded Mendeley from the chart because of its over-proportional counts and coverage. EBS journals are often mentioned in Twitter, News and Blogs showing that journals covering topics of general interest exhibit other social media metrics than journals with a narrower focus.

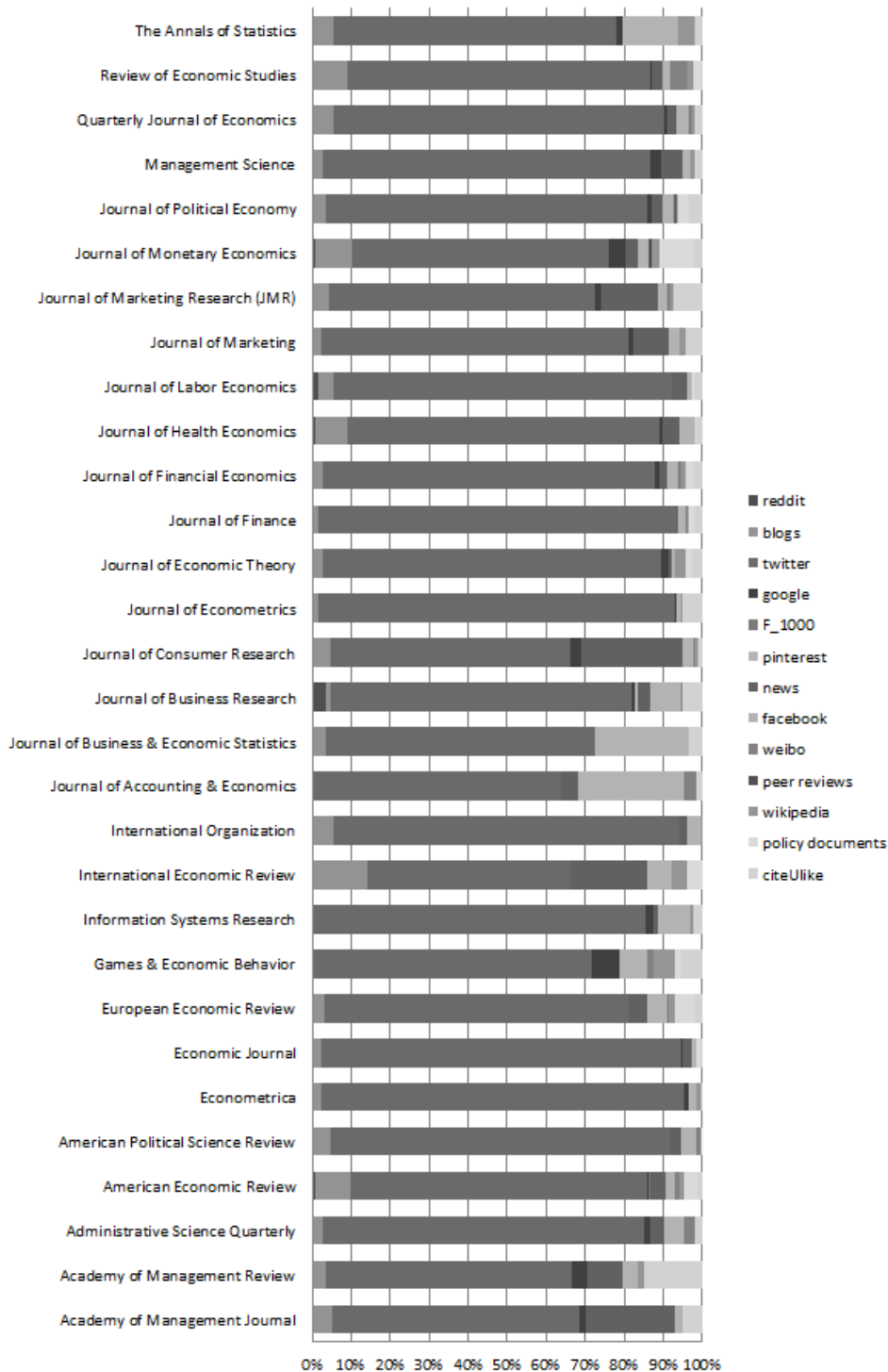
RQ3: Relationship between social media metrics and citation counts

For the publication years 2012-2014, 3,275 articles from 30 EBS journals were found with both citation and altmetric data. The Spearman correlation between citation rates and altmetric scores for 3,275 articles on article level is $r=0.2991$. It indicates a positive but low correlation between these two attributes; however, Spearman correlation for the 3,275 articles between citation and Altmetric Scores on journal level is $r=0.614$. We may speculate here that particular journals are more successful in triggering (or harmonizing) both social media and scientific attention (via citations), but we have to back up this assumption by further investigation.

RQ4: Relationship between IF and Altmetric Score

The correlation between IF and Altmetric Score on journal level is low but positive (Spearman $r=0.314$ and Pearson $p=0.169$) – hence we can conclude that articles from highly cited journals are not receiving substantial attention online. Additionally, articles of highly cited journals are not better covered on social media platforms since no correlation (Pearson $p=0.07$) between the number of DOIs found in Altmetric.com and the IF can be detected.

Figure 2. Social media sources from Altmetric.com on journal level (without Mendeley).

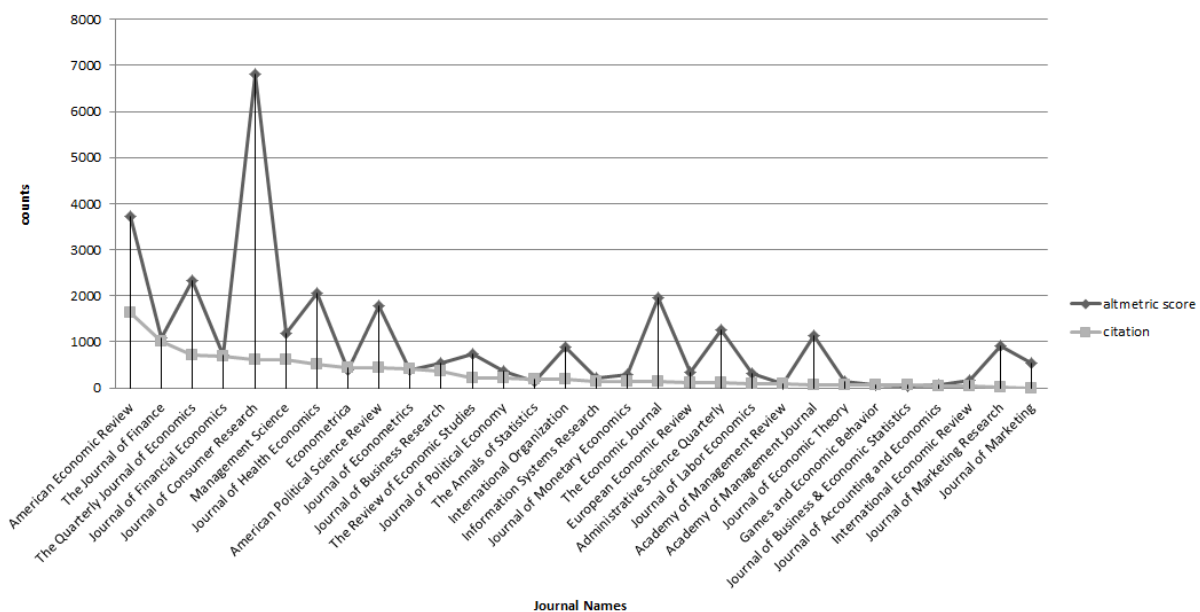


DISCUSSION OF RESULTS

We found that – besides Mendeley – Twitter is the dominating social media platform for EBS journal articles which confirms the results of other studies (Robinson-García et al., 2014). In contrast to other disciplines (Thelwall et al., 2013) blogs are also frequently used for discussion of literature from EBS. As it has been shown (Costas et al., 2015) blogs and tweets have stronger relations with citations and therefore better support identification of highly cited articles. Hence, in EBS indicators derived from engagement with blogs and Twitter may serve as valuable addition to traditional metrics.

The analyses also revealed that for articles in EBS altmetrics data is still rather sparse, although availability increased for more recent articles. However, when considering altmetrics data for real-world application (e.g., in libraries) higher aggregation levels, such as journal level, can well overcome the sparsity of altmetrics data. By doing so, it will be ensured that for every record altmetric information could be displayed which lowers, or even avoids, user frustration.

Figure 3. Comparison of Altmetric Scores and citation counts on journal level.



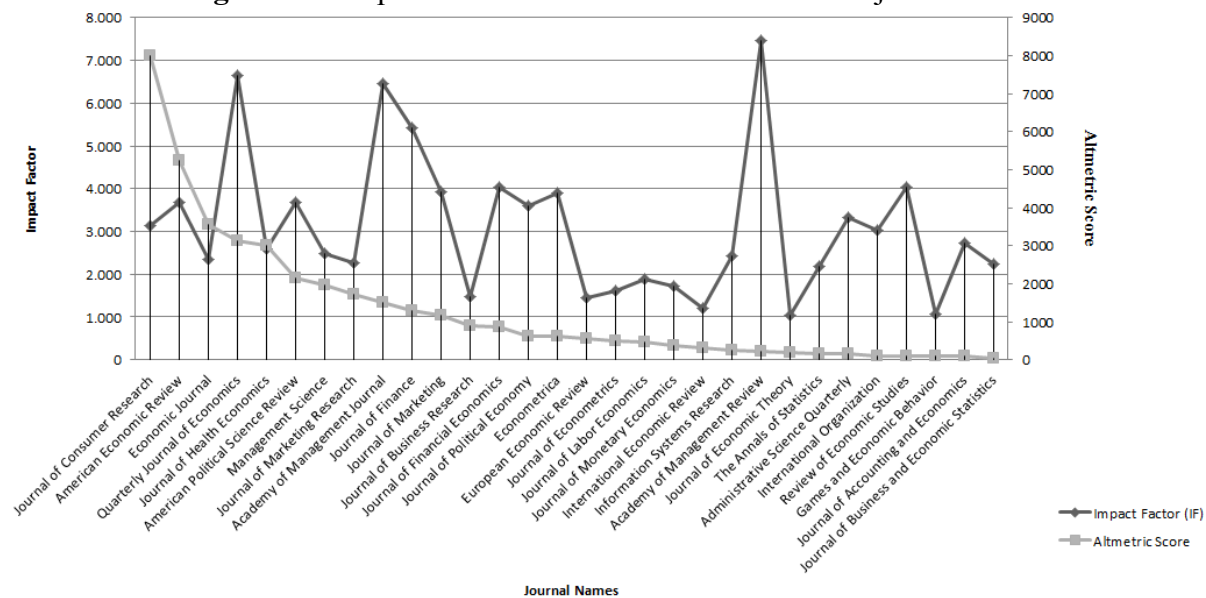
CONCLUSION

We presented an altmetric case study of articles published in the Top 30 journals from Economics and Business Studies by using social media metrics from Altmetric.com. Our results confirm that altmetric information is significantly better present for recent articles. Overall, 38% of articles published in 2012-2014 are represented in Altmetric.com.

The Top 3 most used altmetric sources are Mendeley, Twitter, and News – with Mendeley being the most complete platform for EBS journals (see also Nuredini & Peters, 2015). We could show that Altmetric Scores and citation counts are better correlated on journal level than on article level. On the other hand, the correlation between Altmetric Scores for journals as well as coverage on social media platforms and IFs are low but positive. This shows that 1) articles from highly cited journals do not receive substantial attention online, and 2) altmetrics complement information on the impact of journals provided by traditional indicators.

In order to better understand the relationship of web-based formats of and engagement with scholarly articles, future work will include the analysis of coverage of open access journals from EBS on social media platforms and the expansion of the comparison of Altmetric Scores with citation data by using Google citations.

Figure 4. Comparison of Altmetric scores and IFs on journal level.



REFERENCES

- Costas, R., Zahedi, Z., & Wouters, P. (2015). Do “altmetrics” correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology*, 66(10), 2003–2019. doi: 10.1002/asi.23309
- Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science*, 178, 471–479. doi: 10.1126/science.178.4060.471
- Haustein, S., Peters, I., Sugimoto, C. R., Thelwall, M. & Larivière, V. (2014). Tweeting biomedicine: An analysis of tweets and citations in the biomedical literature. *Journal of the Association for Information Science and Technology*, 65(4), 656–669. doi: 10.1002/asi.23101
- Haustein, S. (2016). Grand challenges in altmetrics: heterogeneity, data quality and dependencies. *Scientometrics*. doi: 10.1007/s11192-016-1910-9
- Nuredini, K., & Peters, I. (2015). Economic and Business Studies Journals and Readership Information from Mendeley. In F. Pehar, C. Schlögl, & C. Wolff, Re: inventing Information Science in the Networked Society, *Proceedings of the 14th International Symposium on Information Science (ISI 2015)*, Zadar, Croatia (pp. 380-392). Glückstadt: Verlag Werner Hülsbusch.
- Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010). Altmetrics: A manifesto. URL: <http://altmetrics.org/manifesto>
- Peters, I., Jobmann, A., Eppelin, A., Hoffmann, C. P., Küne, S., & Wollnik-Korn, G. (2014). Altmetrics for large, multidisciplinary research groups: A case study of the Leibniz

- Association. In *Proceedings of the Libraries in the Digital Age Conference, Zadar, Croatia*. URL: <http://ozk.unizd.hr/proceedings/index.php/lida/article/view/162/138>
- Robinson-Garcia, N., Torres-Salinas, D., Zahedi, Z., & Costas, R. (2014). New data, new possibilities: Exploring the insides of Altmetric.com. *El profesional de la información*, 23(4), 359-366. doi:10.3145/epi.2014.jul.03.
- Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *BMJ*, 314(7079), 498–502. *British Medical Journal*, 314(7079), 498-502.
- Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, CR. (2013) Do Altmetrics Work? Twitter and Ten Other Social Web Services. *PLoS ONE*, 8(5): e64841. doi: 10.1371/journal.pone.0064841
- Zahedi, Z., Costas, R., & Wouters, P. (2015). Do Mendeley readership counts help to filter highly cited WoS publications better than average citation impact of journals (JCS)? In *Proceedings of the 15th International Conference of the International Society for Scientometrics and Informetrics* (pp. 16-25). '
- Wouters, P. & Costas, R. (2012). *Users, Narcissism and Control—Tracking the Impact of Scholarly Publications in the 21 Century*. Utrecht: SURFfoundation. URL: <http://research-acumen.eu/wp-content/uploads/Users-narcissism-and-control.pdf>



Comparative study of Colombian Researchers according to data from Google Scholar, ResearchGate and the National System for Measurement Science (Colciencias)¹

Isidro Aguillo*, Alejandro Uribe-Tirado** and Wilson López López***

*isidro.aguillo@csic.es

Cybermetrics Lab - Scimago Group, IPP-CSIC, C/Albasanz, 26-28. Madrid 28037 (Spain)

**auribe.bibliotecologia.udea@gmail.com, Interamerican School of Information Science - Red ICONOS, University of Antioquia-, Medellín, 1226 (Colombia)

***lopezlopezwilson@gmail.com, Universitas Psychologica Pontifical Xavierian University, Cra. 7 #40-62 Bogotá, (Colombia)

SUMMARY

This paper intends to contextualize the results obtained from the ranking of researchers working at Colombian institutions according to their Google Scholar Citations (GSC) public profiles (1390 with an index equals or larger than 5) with the data from the social network ResearchGate (RG) and the local information provided by Colciencias, the Colombian government agency that publishes a classification of researchers.

The results show significant discrepancies between GSC and RG results with the four categories of the classification provided by Colciencias, suggesting that Colciencias should be reconsider its evaluation criteria including new sources and indicators. As the two sources (GSC, RG) and the indicators (H-index, RG-Index) behave very differently regarding disciplines, Colciencias should also take care of the disciplinary assignments, including developing relative indicators by discipline.

The potential and importance of Internet platforms for visibility and impact (Science 2.0) should be recognized by the Colombian academic and research organizations.

INTRODUCTION

As part of an effort to diversify the sources for bibliometric analysis and to increase the use of author-level metrics in these analysis, the Cybermetrics Lab from CSIC (Spanish Research Council) has started in 2015 the publication of rankings of scientists (by country) according to their Google Scholar Citations (GSC) public profiles, using the h-indexes and total number of citations as ranking criteria. Google Scholar is also providing rankings of researchers by free keywords provided by the authors or by institutions, in both cases using only the decreasing number of citations for arranging the lists. The number of individual and institutional profiles (there are also rankings for journals named as Google Scholar Metrics) is growing fast but the global coverage is still very limited.

With the aim of increasing the number of profiles registered in Latin-American countries, Cybermetrics Lab focused its efforts in this region, including Colombia. For this country an

¹ This work was supported by the EU project ACUMEN and the CSIC (Spain), the University of Antioquia (Colombia) and the Pontifical Xavierian University of Bogotá (Colombia)

updated and enlarged second edition was published in January 2016 with data of 1400 authors (<http://www.webometrics.info/en/node/70>). This list was built selecting the researchers with an h-index larger than 5 from a list of 4863 identified public profiles. A new edition in preparation will evaluate more than 6000 profiles.

This paper intends to contextualize the results obtained from these profiles with the local information provided by Colombian organizations and bibliometric researchers. Thus the profiles of this GSC-based ranking are matched with the data from the social network ResearchGate (RG), and the classification of researchers developed by the local government agency Colciencias.

Colciencias (<http://www.colciencias.gov.co/>) is the Colombian public department in charge of the promotion of research in the academic institutions of the country and the coordination of the National System of Science, Technology, and Innovation SCIENTI (<http://scienti.colciencias.gov.co>) for the monitoring and periodical evaluation of the groups, researchers and scientific journals (<http://publindex.colciencias.gov.co:8084/publindex/>) according to their production with the aim of distribute the resources in a more fairly way, attending the R&D political priorities.

The monitoring and classification process started during the last decade, but it is since the Call 693-2014 in 2015 when a major reorganization of the system took place. The SCIENTI platform provides now a powerful tool for describing the size, structure and performance of the Colombian research effort. Although still contested in many sectors, the data collection that is scheduled to be executed every two years, is becoming the standard source of information for evaluation purposes affecting primarily to individuals and teams and secondarily, but not less important, to their institutions.

The context can be summarized by the classification of the groups (Table 1) and the researchers (Table 2) in categories according to the last 2 Calls. It should be noted that the showed increase is mostly due to the incorporation of many Social Sciences and Humanities teams that finally accepting their registration into the system.

Table 1. Colciencias classification of groups by category

http://www.colciencias.gov.co/sites/default/files/ckeditor_files/resultadosfinales-conv693-2014-consulta.pdf (Checked 4 July 2016)

<http://colciencias.gov.co/sites/default/files/listado-publicacion-resultadosfinales-conv737-gruposinvestigacion-consulta.pdf> (Checked 4 July 2016)

Category	A1	A	B	C	D	Total
GROUPS 693-2014	293	386	869	1543	749	3774
GROUPS 737-2015	408	549	952	1393	610	4458

Table 2. Colciencias classification of researchers by category
<http://colciencias.gov.co/sites/default/files/listado-publicacion-resultadosfinales-conv737-investigadores-consulta.pdf> (Checked 4 July 2016)
http://www.colciencias.gov.co/sites/default/files/ckeditor_files/resultadosfinales-conv693-2014-investigadores-versioncons.pdf (Checked 4 July 2016)

Category	Senior	Associate	Junior	Total
RESEARCHERS 693-2014	1057	2064	5159	8280
RESEARCHERS 737-2015	1218	2767	6057	10042

Perform a triangulation of data possible to identify similarities and differences in these rankings considering that GSC and RG have a tendency to bibliometric, opening to the trends of open access and altmetrics. While Colciencias has been questioned by its tendency to apply more traditional metrics tools, now a discussion is open for a future 2017 expansion of its sources and rating's measurement of researchers, research groups and scientific journals.

METHODOLOGY

For the realization of this study were carried out several phases, which were intended to perform the collection of data from each source (GSC, RG and Scienti Platform Colciencias – CVLAC-), and then begin the integration of such data allow broader interpretations. These phases were:

- Review and debugging of the list of researchers, reported from Webometrics, in its second edition (January 2016), giving a final list of 1379 entries
- Locate the profiles of these investigators at the Scienti (Colciencias) platform and record its researcher classification: *Senior*, *Associate*, *Junior* or *unclassified*
- Identification of the discipline of each one of the researchers from the deepening of the thematic work of each researcher according to the profile information in Google Scholar and scientific production reported
- Classifying the ranked list (by h-index, then by citations number) of researchers in GSC Citations in four quartiles

The list used in this analysis is probably biased, as it consists of mainly the more active authors. Another limitation is related to the disambiguation of names as the profiles are created without a standard identifier like ORCID. Due to this problem it has not been possible to identify all the profiles in the Colciencias and RG databases.

LIMITATIONS

Both web 2.0 sources used in the analysis are set up and maintained mostly by voluntary action of the authors, being ResearchGate by far the most popular and the easiest to build rich individual profiles. Google Scholar Citations is fairly new and most of the users find difficult to locate it and to set up a profile, being unknown for a large number of researchers. The number of active profiles in RG is probably over 8 million, with clearly biased country coverage (Thelwall & Kousha, 2015), whereas the total amount of entries for GSC is still

below one million. The rankings of scientists according GSC have been useful to increase the awareness of this system among LATAM researchers, but the coverage is limited yet, with an informed guess for Colombia that is probably around the 10% of the total population of authors working in the institutions of this country.

Using central measures is being discarded in bibliometric analysis given the skewed distribution (power laws) of the variables. However, a partition approach is becoming popular and recommended, so instead of the full populations with long tails, the analysis focus on the top quartiles or percentiles. Perhaps the threshold chosen here ($h\text{-index} \geq 5$) is pretty low, but it is enough for reducing the sample to the top 25% approximately. On this core, with descriptive purposes, we are using the averages but acknowledging the problems associated.

RESULTS

The results integrate data from the different sources and variables considered, which allows a look, distinct and wide to Colombian investigators, a methodology that can be applicable to other Latin American countries too.

The categories used by Colciencias are almost evenly distributed in our population (Table 1), that is a bit surprising considering that criteria for selection is their high citation performance. A deeper analysis (Table 2) suggests that there are individual cases that can explain the discrepancies, as there are individuals not classified as Seniors with higher h-index than the leaders in that category. It should be taken into account that co-authorship of papers is treated by GSC with full attribution to each one of the authors (100% of authorship), without no distinction of inter-institutional or international cooperation. The GSC country rankings are usually headed by high-energy (particle) physicists working at very large organizations like the CERN. For example, the papers describing the recent discovery of the Higgs' boson are signed usually for more than 1000 authors (databases mention more than 600 papers over that figure with at least one with over 5000 names). There are Colombian authors with GSC profiles in these papers.

Table 3. Distribution by CVLAC categories of the Researchers

	Categories	Researchers	%	Totals
Excluded	<i>No presence</i>	99		202
	<i>Unmatched</i>	84		
	<i>Incomplete</i>	19		
CVLAC Categories	<i>No category</i>	270	22,9%	1177
	<i>Junior</i>	344	29,2%	
	<i>Associate</i>	250	21,2%	
	<i>Senior</i>	313	26,6%	
TOTAL				1379

Table 4. GSC indicators for the CVLAC categories

CVLAC Categories	AVERAGE	MAXIMUM		
		H- CITATIONS INDEX	H- CITATIONS INDEX	
<i>No category</i>	450	8,6	21144	75
<i>Junior</i>	406	7,9	43981	93
<i>Associate</i>	340	7,8	8594	22
<i>Senior</i>	947	13	16371	64

The Tables 3 & 4 explores the population according to its distribution by quartiles (equal groups from the ranking). As expected the number of Seniors in the first quartile is far larger than the rest of the groups and far lower regarding the fourth quartile, just the inverse that shows the Juniors and Associates distribution. A similar pattern arises when the number of citations received is considered.

Table 5. Number of profiles by CVLAC categories and GSC quartiles

GS PROFILES / QUARTILES					
CATEGORY	QUARTIL 1	QUARTIL 2	QUARTIL 3	QUARTIL 4	
<i>Senior</i>	151	82	57	23	
<i>Associate</i>	35	68	65	82	
<i>Junior</i>	46	74	110	114	
<i>Unclassified</i>	49	60	67	94	1177
<i>NE</i>	64	61	46	31	
Total					1379

Table 6. Number of citations in GSC distributed by quartiles and CVLAC categories

CITATIONS: GS					
CATEGORY	QUARTIL 1	QUARTIL 2	QUARTIL 3	QUARTIL 4	
<i>Senior</i>	246014	36603	11376	2287	
<i>Associate</i>	36873	28227	11767	8109	
<i>Junior</i>	90546	20858	17930	10415	
<i>Unclassified</i>	78309	19440	13100	10736	
<i>NE</i>	86464	21892	8051	3954	
					762951

In order to check if the patterns are related with the disciplines, we calculate the average number from both h-index and citations from Google Scholar according to the Schools where the authors are affiliated (Table 5). Sciences and technologies are well represented, with Medicine and Health Sciences (Odontology, Nursing, Pharmacy) as the top disciplines.

Table 7. Average values of h-index and number of citations by discipline (School/ Faculty affiliation)

DISCIPLINE	AVERAGE	
	H-INDEX	CITATIONS
Medicine	13	196
Agronomy	11	573
Science	10	681
Health Sciences	10	709
Environmental and Rural Studies	10	609
Technologies	10	313
Geosciences	10	460
Social Sciences	9	433
Economic and Administrative Sciences	9	363
Engineering	8	388
Law and Political Science	8	340
Education	8	298
Psychology	8	291
Environmental Sciences	8	291
Odontology	8	1276
Communication and Language	7	218
Nursing	7	164
Arts	7	287
Architecture and Design	7	144
Pharmaceutical Sciences	7	224
Technology Center for Academia	6	133
CEDEX	5	125
Editor	5	76
Average Overall Total	9	553

These results are consistent with the traditional bibliometric patterns as in fact Google Scholar is a very large, but standard, citation database. An alternative source of metrics is the academic social network ResearchGate (<http://www.researchgate.net>) that shows similar results to the ones already obtained from Google Scholar: Seniors are consistently reaching higher values in all the RG indicators, although it looks there are special cases of ‘rising stars’ among the Juniors (Tables 6 &7). Hyper-authorship can explain again a few of these results, but RG indicators takes into account not only the citations but also the social activity of the authors with other members of the network. Juniors can be more free of managerial duties than their seniors’ counterparts and can devote more time for attending requests or questions from colleagues at RG. As it can be expected there is also a generation gap regarding the use of electronic media that should be considered as usually the academic age of the juniors is lower than of the rest of the faculty members.

Table 8. RG indicators for the CVLAC categories

CVLAC Categories	AVERAGE		MAXIMUM	
	CITATIONS	RG- INDEX	CITATIONS	RG- INDEX
No category	262,31	16.01	7190	50,88
Junior	171,16	14,08	2533	64,35
Associate	164,76	14,42	1177	31,25
Senior	521,68	20,41	11239	48,26

Table 9. Average values of h-index and RG-score by discipline (School/ Faculty affiliation)

DISCIPLINE	AVERAGE	
	RG SCORE	H- INDEX
Technology Center for Academia	24,17	13
Pharmaceutical Sciences	20,13	10
Communication and Language	26,27	10
Odontology	24,10	10
Health Sciences	9,57	10
Psychology	17,77	10
Education	10,32	9
Law and Political Science	10,55	9
Technologies	15,52	8
Environmental and Rural Studies	5,58	8
Engineering	5,97	8
Environmental Sciences	12,19	8
Social Sciences	15,59	8
Science	16,88	8
Nursing	8,76	7
Medicine	6,84	7
Science	ND	7
Environmental Sciences	12,09	7
Editor	16,73	7
Economic and Administrative Sciences	8,62	6
Health Sciences	ND	5
Geosciences	ND	5
Average Overall Total	16,93	9

Finally, we analyzed the results using the data for the most representative institutions, all of them (public and private) universities. Using the group data from the last two Colciencias Calls, the Observatorio de la Universidad Colombiana (<http://www.universidad.edu.co/>) developed a points system granting different weights to the groups according to their performance in the Colciencias classification. For the full list of the Call 737-2015 (the most

recent one) the correlation with the total number of citations obtained in the GSC ranking is pretty high ($R^2=0.708^{**}$) for the 50 universities best ranked. The Table 10 shows the comparative data for the top 25 universities.

Table 10. Comparative performance of 25 Colombian Universities according to the total number of citations in the selected GSC profiles and the points attributed according to the number and classification of their research groups provided by

Colciencias Call 693-2014

(http://www.universidad.edu.co/index.php?option=com_content&view=article&id=5715:2015-05-10-12-04-18)

Colciencias Call 737-2015

(http://www.universidad.edu.co/index.php?option=com_content&view=article&id=6954:2016-06-04-16-59-15)

UNIVERSITY	CITATIONS	POINTS	POINTS
	GSC	2014	2015
Universidad Nacional de Colombia	138789	395	447.6
Universidad de Antioquia	85369	201.4	215.6
Universidad del Valle	37232	112	135.8
Universidad de Los Andes Colombia	118012	117.8	128.6
Pontificia Universidad Javeriana	66022	91	99
Universidad Distrital Francisco José de Caldas	5968	55.8	77
Universidad Industrial de Santander	20680	70.4	72.2
Universidad de Cartagena Colombia	16252	52.4	65.2
Universidad Pedagógica y Tecnológica de Colombia	1818	44.4	61.8
Universidad Pontificia Bolivariana	6501	55.2	60.2
Universidad Libre	505	50.4	58.4
Universidad Tecnológica de Pereira	8493	47.4	57.2
Universidad de Caldas	1925	33.8	46.2
Universidad del Cauca	6504	32.2	42.4
Universidad del Atlántico	1402	33.6	42.2
Universidad del Norte	18158	34.6	41.8
Universidad Santo Tomás	4329	31.4	41.2
Universidad de La Sabana	2014	34.8	38.6
Universidad Militar Nueva Granada	1817	33.2	38.2
Universidad Pedagógica Nacional	971	27.2	38
Universidad EAFIT	8669	32.8	36.8
Universidad de Nariño	362	28.4	35.8
Universidad del Rosario	31989	32.4	34
Universidad Cooperativa de Colombia	1528	30.6	33.8
Universidad del Tolima	3018	28	33.6

It looks that the institutional performance based on the research classification of the groups is similar to the one provided by the more transparent GSC system. Strikingly the classification of individuals does not correlate so well, even taking into account the large coverage of Colciencias CV system and the very inclusive Google Scholar database, far larger than the WoS or Scopus ones.

CONCLUSIONS AND PERSPECTIVES

If we consider the comparative tables between GS and RG with the four classifications CVLAC Colciencias, is shown that a significant percentage of researchers who are as *Unclassified* or *Juniors* however showed an important web visibility and impact. On the other side, an important percentage of researchers classified as *Seniors* or *Associates*, have low values in GS and RG, even considering that their indicators of scientific production, as reported in the CVLAC, are very high. This implies that the sources used for the classification of Colciencias should be modernized.

Although more information is needed, the two sources (GSC, RG) and the indicators (H-index, RG-Index) are behaving very differently regarding disciplines. Again that means Colciencias should make careful disciplinary assignments for comparative purposes, including developing relative indicators by discipline.

The potential and importance of Internet platforms for visibility and impact (Science 2.0) is already starting to be recognized by some Colombian universities (Uribe-Tirado, 2015), but Colciencias still have a relevant role in training/convincing on these issues to researchers, as it was already indicated Kuchner (2012): "Being a good scientist is half science and half marketing".

REFERENCES

- Kuchner, Mark J. (2012). *Marketing for Scientist*. Washintong: IslandPress, 236 p.
- Mas-Bleda, A., Thelwall, M., Kousha, K., & Aguillo, I.F. (2014). Do highly cited researchers successfully use the social web? *Scientometrics*, 101(1), 337–356.
- Orduña-Malea, E. & Aguillo, I.F. (2014). *Cibermetría: midiendo el espacio red*. Barcelona: EPI-UOC, 194 p.
- Ortega, J.L. & I.F. Aguillo (2014). Microsoft Academic Search and Google Scholar Citations: A comparative analysis of author profiles. *Journal of the American Society for Information Science and Technology*, 65 (6), pp. 1149–1156
- Thelwall, Mike; Kousha, Kayvan (2015). ResearchGate: Disseminating, communicating, and measuring Scholarship?. *Journal of the Association for Information Science and Technology*, 66 (5), pp 876–889. DOI: 10.1002/asi.23236
- Uribe-Tirado, A. (2015). *Percepciones, políticas y formación hacia la visibilidad académica y científica entre profesores-investigadores de la Universidad de Antioquia*. In: 15 Conferencia Internacional BIREDIAL-ISTEC, Barranquilla-Colombia, Universidad del Norte, 17 al 21 de noviembre. Retrieved March 2, 2016 from:<http://eprints.rclis.org/28499/>



Next-Generation Metrics: Responsible Metrics & Evaluation for Open Science

James Wilsdon^{*}, Judit Bar-Ilan^{**}, Isabella Peters⁺ and Paul Wouters⁺⁺

^{*}j.wilsdon@sheffield.ac.uk

Department of Politics, University of Sheffield, 241 Glossop Road, Sheffield S10 2GW (UK)

^{**}Judit.Bar-Ilan@biu.ac.il

Department of Information Science, Bar-Ilan University, Ramat Gan, 5290002 (Israel)

⁺i.peters@zbw.eu

Web Science, ZBW Leibniz Information Centre for Economics & Kiel University, Duesternbrooker Weg 120, Kiel, 24105 (Germany)

⁺⁺p.f.wouters@cwts.leidenuniv.nl

Centre for Science and Technology Studies (CWTS), Leiden University, P.O. Box 905 2300 AX Leiden (The Netherlands)

INTRODUCTION

Metrics evoke a mixed reaction from the research community. A commitment to using data to inform decisions makes some enthusiastic about the prospect of granular, real-time analysis of research and its wider impacts. Yet we only have to look at the blunt use of metrics such as journal impact factors, h-indices and grant income targets, to be reminded of the pitfalls.

Some of the most precious qualities of academic culture resist simple quantification, and individual indicators often struggle to do justice to the richness and plurality of research. Too often, poorly designed evaluation criteria are “dominating minds, distorting behaviour and determining careers (Lawrence, 2007).”

Metrics hold real power: they are constitutive of values, identities and livelihoods. How to exercise that power to more positive ends has been the focus of several recent and complementary initiatives, including the San Francisco Declaration on Research Assessment (DORA¹), the Leiden Manifesto² and The Metric Tide³ (a UK government review of the role of metrics in research management and assessment).

Building on these initiatives, the European Commission, under its new Open Science Policy Platform⁴, is now looking to develop a framework for responsible metrics for research management and evaluation, which can be incorporated into the successor framework to Horizon 2020.

¹ <http://www.ascb.org/dora/>

² <http://www.leidenmanifesto.org/>

³ http://www.hefce.ac.uk/pubs/rereports/Year/2015/metrictide/Title_104463_en.html

⁴ <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-policy-platform>

A European Commission expert group has been formed to lead this work, and will report to Commissioner Moedas and the Open Science Policy Platform by the end of 2016.⁵

This panel at STI 2016 will consist of the chair (Wilsdon) and three members (Bar-Ilan, Peters and Wouters) of the new EC expert group, who will discuss the background, goals and current state of their assignment, and invite feedback and input from the scientometrics and science/innovation policy community at STI 2016.

Topics and Purpose of Panel

Issues to be addressed will include:

- Categorisation and review of new and emerging metrics and their relationship to more established scientometrics;
- The prospects for ‘responsible metrics’ (cf. Wilsdon et al. 2015; Hicks et al. 2015) to advance more open, diverse, transparent and robust research;
- Landscape and comparative analysis of what’s going on in this area across the EU and internationally;
- Implications of different metrics and indicators for equality and diversity (including diversity of discipline/field);
- Gaming, strategic responses and manipulation of particular indicators;
- Use of metrics and indicators to inhibit, encourage and recognize interdisciplinary work;
- Metrics and indicators to support and measure diverse pathways to impact;
- The implications of emerging social networks for scientists (ResearchGate, Academia, Mendeley); research information and management systems (e.g. Pure, SciVal) and citation profiles (Google Scholar) for open science;
- Possible use of metrics and altmetrics in the next framework programme (including the ERC), in support of open science.

More specifically, the activities which will contribute to the understanding of these issues and which have to be carried out by the expert group are:

- To assess the changing role of metrics in measuring and evaluating the qualities and impacts of research;
- To consider how metrics can be developed and used responsibly to support the development of open science;
- To engage stakeholders and identify their needs in terms of research metrics, and associated indicators of qualities and impacts;
- To consider the implications of existing and newer metrics for diversity and equality; interdisciplinarity; research cultures; gaming and other strategic responses.
- To examine the implications of emerging social networks for scientists; research information and management systems; and citation profiles;
- To explore possible uses of metrics and altmetrics in tracking impacts, research actions and deliverables under Horizon 2020, and within the next framework programme (including the ERC), in support of open science;
- To consider the data infrastructure required to underpin robust, responsible, transparent and interoperable uses of metrics and altmetrics in support of open science.

⁵ http://ec.europa.eu/research/openscience/index.cfm?pg=altmetrics_eg

Activities and Agenda

In this 90-minutes session, three members of the EC expert group (Bar-Ilan, Peters and Wouters) will each speak for ten minutes about aspects of the group's work, preliminary results from a small stakeholders' workshop, and from a call for evidence. Wilsdon will chair and moderate the session, which will include a significant amount of time for debate and discussion (e.g. via a fish bowl discussion). The panellists will outline key questions for discussion and ask for additional input, which will feed into the final report of the EC group.

Anticipated participants

It is hoped that the participants of STI 2016 will engage in a lively discussion and contribute to the expert group by broadening its scope, and offering insights from a wide range of stakeholder perspectives, including: universities; researchers; learned societies; publishers; developers and practitioners of metrics and altmetrics.

PANELISTS:

James Wilsdon is Professor of Research Policy and Director of Impact and Engagement in the Faculty of Social Sciences at the University of Sheffield. Since 2013, he has been Chair of the UK's Campaign for Social Science, and he recently chaired an independent review of the role of metrics in the management of the UK's research system, which published its final report *The Metric Tide* in July 2015. Previously, he worked as Professor of Science and Democracy at SPRU, University of Sussex (2011-2015), Director of Science Policy at the Royal Society (2008-2011), Head of Science and Innovation at Demos (2001-2008), Senior Research Fellow at Lancaster University's Institute for Advanced Studies (2006-2008) and Senior Policy Adviser at Forum for the Future (1997-2001). James is one of the editors of the Guardian's 'Political Science' blog, on science, research and innovation policy, and an Associate Editor of the open access journal *Palgrave Communications*. In 2015, he was elected a Fellow of the Academy of Social Sciences, the UK's national academy of academics, learned societies and practitioners in the social sciences. He is on twitter @jameswilsdon.

Judit Bar-Ilan is Professor at the Department of Information Science at Bar-Ilan University. She holds the Chair of the Library Committee and headed the Department of Information Science from 2008-2012. She is the academic head of the Israeli Consortium for Digital Information Services and she earned her Ph. D in computer science from The Hebrew University of Jerusalem in 1990.

Isabella Peters is Professor of Web Science at ZBW Leibniz Information Center for Economics and Kiel University. She has been Professor of Web Science at ZBW Leibniz Information Centre for Economics and Chair of the Web Science research group at Kiel University since 2013. She received her PhD in Information Science at the Heinrich Heine University in Düsseldorf. Her research focusses on user-generated content and its potential in knowledge representation and information retrieval as well as on scholarly communication on the social web, e.g. altmetrics. Professor Peters is active in the Association for Information Science and Technology (in particular European Chapter and SIGMetrics), in the LIBER Working Group on Metrics as well as in the Leibniz Research Alliance Science 2.0.

Paul Wouters is Director of the Centre for Science and Technology Studies and Professor in Scientometrics. He has published on the history of the Science Citation Index, on scientometrics and on the way the criteria of scientific quality have been changed by citation analysis. He has also studied the role of information and information technologies in the creation of new scientific and scholarly knowledge. In this area, he was appointed as leader of two research programs by the Royal Netherlands Academy of Arts and Sciences: *Networked Research and Digital Information (Nerdi)* (2000 - 2005) and *The Virtual Knowledge Studio for the Humanities and Social Sciences (VKS)* (2005 - 2010). He is a member of the editorial board of *Social Studies of Science*, *Journal of the Association of Information Science and Technology*, and *Cybermetrics* and sits on various advisory boards of international programs and projects. Recently, he was co-author of [The Metric Tide](#) and [The Leiden Manifesto for Research Metrics](#) on principles for the use of metrics in research assessment.

REFERENCES

- Lawrence, P.A. (2007) The mismeasurement of science. *Current Biology*, 17(15), 583–R585.
- Hicks, Diana, Paul Wouters, Ludo Waltman, Sarah de Rijcke, and Ismael Rafols. 2015. “The Leiden Manifesto for Research Metrics.” *Nature* 520: 429–31. doi:10.1038/520429a.
- Wilsdon, J. et al. 2015. *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. doi:10.13140/RG.2.1.4929.1363.



A Systematic Identification of Scientists on Twitter¹

Qing Ke*, Yong-Yeol Ahn* and Cassidy R. Sugimoto*

*qke@indiana.edu; yyahn@indiana.edu; sugimoto@indiana.edu
School of Informatics and Computing, Indiana University Bloomington, 47408 (US)

ABSTRACT

There is an increasing use of Twitter and other social media to estimate the broader social impacts of scholarship. However, without systematic understanding of the entities that participate in conversations about science, efforts to translate altmetrics into impact indicators may produce highly misleading results. Here we present a systematic approach to identifying scientists on Twitter.

INTRODUCTION

Twitter and other social media channels are increasingly used by various stakeholder groups in science. For instance, scientists disseminate and discuss research results (Darling & Shiffman, 2013; Faulkes, 2014; Woolston, 2015); practitioners engage in journal clubs and other community-based activities (Lulic & Kovic, 2013); and health policy makers initiate direct dialogues with scientists (Kapp, Hensel & Schnoring, 2015). Quantitative measures of scholarly activities on social media—often called altmetrics (Priem, Piwowar & Hemminger, 2012)—can now be studied at scale, given the transparency and availability of APIs on several platforms, most notably Twitter. Altmetrics have been compared to citation-based metrics in extant literature, showing low correlation. Such low correlation has been argued to speak to an alternative measure of impact, particularly the broader impact on the society (Bornmann, 2014), given that social media provides an open platform where people with diverse backgrounds can engage in direct conversations. However, this argument has not been empirically tested, and many related questions remain unanswered, impeding our further understanding of the usage of altmetrics for measuring the broader impact of scientific papers.

A necessary first step towards resolving these challenges is to examine *who* is engaging in scientific discourse and generating relevant activity traces on social media (the dominant source for altmetrics). The underlying argument is that altmetrics broadens our understanding of scientific impact, by demonstrating that altmetric activity is generated not by scientists, but by “the public.” In order to empirically validate these claims, we need to be able to identify scientists and non-scientists. Previous attempts to identify scientists have suffered from the limitations of focusing on single pre-selected disciplines (Haustein, Bowman & Holmberg, 2014; Hadgu & Jaschke, 2014; Holmberg, Bowman, Haustein & Peters, 2014) and/or a small number of scientists (Haustein, Bowman & Holmberg, 2014; Haustein, Bowman & Holmberg, 2014; Holmberg & Thelwall, 2014). Moreover, previous studies use purposive sampling: first identifying a few scientists (often highly cited people) in another source and then identifying them on a social media platform. This process relies on external bibliographic databases and is bound by the parameters of other indicators (e.g., number of citations). Here, we propose a method to organically identify scientists on one specific platform (i.e., Twitter) without using external bibliographic databases, and present a descriptive analysis of the

¹ CRS is supported by the Alfred P. Sloan Foundation Grant #G-2014-3-25

largest set of scientists on Twitter ever collected. Our method can be easily modified and adapted to identify other groups of stakeholders and will serve as a basic building block to study scholarly communication and the broader impact of science.

METHOD

Our method consists of three steps: (1) compiling a list of scientist titles, (2) curating initial scientist users, and (3) expanding them through Twitter's lists to discover more.

Scientist Occupations

One of the first hurdles to overcome is the operationalization of a "scientist." To address this, we first create a list of occupation titles of scientists. We use the 2010 Standard Occupational Classification (SOC) system (<http://www.bls.gov/soc/>) released by the U.S. Bureau of Labor Statistics. The SOC provides not only practical and authoritative guidance about whether some particular occupations should be considered as scientists, but also official statistics (e.g., total employment of social scientists) that can be used for demographic analysis. SOC is a hierarchical system that classifies workers into 23 major occupational groups, among which we focus on two categories: (1) Computer and Mathematical Occupations, and (2) Life, Physical, and Social Science Occupations. 28 scientist occupation titles are compiled from the two categories. We augment this with Wikipedia, to make the list more comprehensive and contemporary (http://en.wikipedia.org/wiki/Scientist#By_field). For each title, we consider its singular form and the core disciplinary term (e.g., for "clinical psychologists," we also consider "clinical psychologist," "psychologists," and "psychologist"). We add two generic titles to the list: "scientists" and "researchers." In total, 322 scientist titles are generated in this method.

List-based Identification of Scientists

We identify scientists on Twitter by examining the information that other users provide about a particular user, through Twitter lists. A Twitter list contains a set of Twitter users. To create a list, one provides a name and optional description. Although the purpose of lists is to help users organize their subscriptions, the names and descriptions of lists are helpful in identifying attributes, particularly expertise, of users in the lists (Sharma et al., 2012). By extracting all the descriptions from lists associated with users we can effectively "crowdsource" and answer whether a given Twitter user is a scientist or not as well as what kinds of scientists the user is.

In principle, we could use Twitter's memberships API, for each user, to get all the lists containing this user, and then infer whether this user is a scientist by analyzing the names and descriptions of these lists. However, this method is too time-consuming and almost infeasible because (1) a large fraction of users may not be scientists, (2) the distribution of number of listed times for users is right-skewed, and (3) Twitter API has rate limits.

We instead employ a list-based snowball sampling method, starting from a given initial set of scientists and expanding to discover more. A similar method was previously introduced in (Wu et al., 2011). We introduce two major improvements: first, we use a more authoritative and comprehensive list of scientist titles; and second, we use a systematic way to curate the initial user set.

Getting initial seed users. The most frequently used words in the names and descriptions of the lists containing a user are considered as the “attributes” of the user. We leverage already collected attributes of users from the website <http://twitter-app.mpi-sws.org/who-is-who>, which takes the screen name of a Twitter user as input and returns a word cloud for the given user with font sizes of words encoding the frequency of their appearance in list names and descriptions. Note that in the website, attributes are only available for those users who are included in at least 10 lists.

With the Twitter who-is-who interface, we crawl the attributes of 2,436,889 users. In doing so, we first collect 285,760,507 unique users by scanning a Twitter Gardenhose dataset, which contains about 10% of all public tweets from January 2013 to June 2014. The number of users is comparable to the number reported in a previous large-scale Twitter study, and the set of users covers any account that tweeted at least once and at least one of these tweets is included in Gardenhose during the period. We then filter out those users who were listed less than eight times and query all the remaining users on the who-is-who website, finally obtaining attributes of about 2.4M users.

From the 2.4M users, we obtain seed users who are most likely to be scientists. As these seeds will be used for expansion, we prefer precision to recall. We thus adopt stringent criteria to filter out non-seed users. Specifically, we first disregard the least important attributes of each user and then keep those users whose attributes contain the attribute “science” and at least one scientist title complied before. The obtained initial set has 8,545 users, and we use them as initial seeds for snowball sampling.

Snowball sampling. We use snowball sampling (breadth-first search) on the bipartite network of Twitter users and lists. For each public user in queue, we get all the lists in which the user appears using the Twitter memberships API. Among those lists, the lists whose name contains at least one scientist title are kept and all of their members are obtained. Those who have not been visited are put into the queue. These steps are repeated until the queue becomes empty, which completes the sampling process. Note that during the sampling, we only consider users whose names contain spaces. This process removes many organizations (non-person) and the users who are more anonymous. It may also drop many users with non-English names.

From the sampling procedure, we arrive at 110,708 users appearing in 4,920 lists with scientist titles. To increase the precision of our method, the final dataset contains those users whose profile descriptions also contain scientist titles. A total number of 45,867 users are found. Table 1, which shows the top 30 scientists based on number of lists whose names contain scientist titles, suggests that our sampling method can identify scientists in diverse disciplines.

Table 1. Top 30 scientists based on the number of Twitter lists whose names contain scientist titles.

Name	Discipline	Lists	Name	Discipline	Lists
Michio Kaku	Physicist	190	Sam Harris	Neuroscientist	77
Richard Dawkins	Biologist	189	Barry Eichengreen	Economist	75
Sean Carroll	Physicist	141	Brian Greene	Physicist	75
J. Bradford DeLong	Economist	136	Carolyn Porco	Planetary scientist	74
Steven Pinker	Cognitive scientist	135	Danah Boyd	Social media scholar	69
Neil deGrasse Tyson	Astrophysicist	133	Katherine Mack	Astrophysicist	65
Jonathan Eisen	Biologist	102	Richard H. Thaler	Economist	65
Tim Harford	Economist	102	Miles Kimball	Economist	63
Paul Zachary Myers	Biologist	100	Lisa Randall	Physicist	60
Lawrence M. Krauss	Physicist	96	Mike Brown	Astronomer	59
Dan Ariely	Economist	93	Robert J Shiller	Economist	59
Karen James	Biologist	85	Hilary Mason	Data scientist	59
Jim Al-Khalili	Physicist	84	Greg Mankiw	Economist	58
Richard Wiseman	Psychologist	77	J. Craig Venter	Life scientist	57
Betsey Stevenson	Economist	77	Andrew David Thaler	Ecologist	57

WORK-IN-PROGRESS

This paper provides a methodological contribution that is necessary for those conducting large-scale altmetric research. In order to assess impact, one must have a better understanding of the audience that generates these metrics. This paper provides a method for accomplishing this for one particular stakeholder group; however, more information is necessary to inform our understanding of this group. For example, these data could be described according to demographics and content of tweets. Furthermore, the network information from these data could be leveraged both to identify additional scientists on the platform and to provide insights into the social structure of Twitter. These data will be explored and presented as part of this work-in-progress.

REFERENCES

- Bornmann, L. (2014). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *J Informetr*, 8, 895-903.
- Darling, E.S., Shiffman, D., Cote I.M. & Drew J.A. (2013). The role of Twitter in the life cycle of a scientific publication. *Ideas in Ecology and Evolution*, 6, 32-43.
- Faulkes, Z. (2014). The Vacuum Shouts Back: Postpublication Peer Review on Social Media. *Neuron*, 82, 258-260.
- Hadgu, A.T., Jaschke, R. (2014). Identifying and Analyzing Researchers on Twitter. In: Proc. of the 2014 ACM Conference on Web Science, p. 23-32.
- Haustein, S., Bowman, T.D., Holmberg, K., Peters I. & Lariviere, V. (2014). Astrophysicists on Twitter. *AJIM*, 66, 279-296.
- Holmberg, K., Bowman, T.D., Haustein, S. & Peters, I. (2014). Astrophysicists' Conversational Connections on Twitter. *PLoS ONE*, 9, e106086.
- Holmberg, K & Thelwall, M. (2014). Disciplinary differences in Twitter scholarly communication. *Scientometrics*, 101, 1027-1042.
- Kapp, J.M., Hensel, B. & Schnoring, K.T. (2015). Is Twitter a forum for disseminating research to health policy makers? *Annals of Epidemiology*, 25, 883-887.
- Lulic, I. & Kovic, I. (2013). Analysis of emergency physicians' Twitter accounts. *Emergency Medicine Journal*, 30, 371-376.
- Priem, J., Piwowar, H.A., Hemminger, B.M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. arXiv:1203.4745.
- Sharma, N.K., Ghosh, S., Benevenuto F., Ganguly, N. & Gummadi, K. (2012). Inferring Who-is-who in the Twitter Social Network. In: Proc. of the 2012 ACM Workshop on Online Social Networks, p. 55-60.
- Woolston, C. (2015). Potential flaws in genomics paper scrutinized on Twitter. *Nature*, 521-397.
- Wu S., Hofman J.M., Mason W.A., Watts D.J. (2011). Who says what to whom on Twitter. In: Proc. of the 20th International Conference on World Wide Web, p. 705-714.



Do Mendeley reader counts reflect the scholarly impact of conference papers? An investigation of Computer Science and Engineering fields

Kuku Joseph Aduku, Mike Thelwall, Kayvan Kousha.

K.J.Aduku@wlv.ac.uk; M.Thelwall@wlv.ac.uk; K.Kousha@wlv.ac.uk

*Statistical Cybermetrics Research Group, School of Mathematics and Computer Science,
University of Wolverhampton, Wulfruna Street, Wolverhampton WV1 1LY, UK*

ABSTRACT

Counts of Mendeley readers may give useful evidence about the impact of research. Although several studies have indicated that there are significant positive correlations between counts of Mendeley readers and citation counts for journal articles, it is not known how the pattern of association may vary between journal articles and conference papers. To fill this gap, Mendeley readership data and Scopus citation counts were extracted for both journal articles and conference papers published in 2011 in four fields for which conferences are important; Computer Science Applications, Computer Software, Building & Construction Engineering and Industrial & Manufacturing Engineering. Mendeley readership counts were found to correlate moderately with citation counts for both journal articles and conference papers in Computer Science Applications and Computer Software. Nevertheless, the correlations were much lower between Mendeley readers and citation counts for conference papers than for journal articles in Building & Construction Engineering and Industrial & Manufacturing Engineering. Hence, there seems to be disciplinary differences in the usefulness of Mendeley readership counts as impact indicators for conference papers, even between fields for which conferences are important.

Keywords: Mendeley readers; Citation counts; Journal articles; Conference papers; Correlations.

INTRODUCTION

Mendeley readership counts are promising indicators for scholarly impact because they appear much earlier than citations because they are as a consequence of not being affected by publication delays. For example, an article may be registered in Mendeley on the day that it is published. Mendeley readership counts can also reveal the disciplines and nationalities of authors, giving more specific impact evidence (Thelwall & Sud, 2015). Investigations of Mendeley readership counts so far have focused on either journal articles or books, but conference papers are valuable in some engineering-related fields and so it is important to assess whether they could also be applied to conference papers as a usage indicator.

Although there are many limitations with using citation counts in formal and informal research evaluation as scholarly impact indicators, they are more robust than indicators

derived from the web because these can easily be manipulated, making them unsafe for most formal evaluations (Wouters & Costas, 2012). Several years ago, Mendeley had about 2.4 million users who uploaded over 420 million documents across disciplines, ranging from life science to math to the arts and humanities (Gunn, 2013). Although Mendeley could be spammed, its large user base and positive results from previous analyses with it (see below) suggest that it does not currently suffer from a substantial amount of spam.

Papers presented at conferences in many fields are seen as a stage towards the creation of journal articles (Drott, 1995). Nevertheless, in some fields conference papers are valued for being timelier, more cutting-edge and more cited than journal articles (Goodrum, McCain, Lawrence, & Giles, 2001) and can be either regarded as the main outputs of research or broadly comparable to journal articles as research outputs.

The gap that this research tries to fill is to discover whether the impact of conference papers is reflected in their Mendeley readership counts in engineering-related fields in which they are important. Although several studies have found correlations between Mendeley readership counts and citation counts (Li, Thelwall, & Guistini, 2011; Bar-Ilan, 2012; Mohammadi & Thelwall, 2014), the extent to which Mendeley readership counts capture the impact of conference papers is unknown. The current study fills this gap by comparing Mendeley readership counts and citation counts for both journal articles and conference papers in four engineering-related Scopus subject categories: Computer Science Applications, Computer Software, Building & Construction Engineering and Industrial & Manufacturing Engineering.

ALTMETRICS

The term altmetrics (alternative metrics) refers to academic indicators derived from social web data. Altmetrics rely on real-time data and interactions, can be quantified, and measured immediately (Galloway, Pease & Rauh, 2013). Existing altmetrics have used a variety of data sources including article downloads (Bollen, Vandel Sompel, & Rodriguez, 2008) views and saves, as well as tweets, blogs, bookmarking sites and wikis are used by scholars to communicate different kinds of research impact (Cronin, 2013a) and can be used to capture different types of non-scientific research impact (Bornmann, 2014). Although altmetrics have the potential to identify scholarly impacts, there is need for more study to substantiate this claim.

The main, but not only, way to assess altmetrics is through correlation tests (Sud, & Thelwall, 2014). One altmetric, tweet counts, might not be suitable for correlation tests, based on its increasing uptake resulting in newer articles having higher tweet counts than the older articles. In contrast, citations take time to accrue and so tests for correlations with altmetrics indicator might be biased towards negative correlations (Thelwall, Haustein, Larivière, & Sugimoto, 2013). There is no evidence that this issue also applies to Mendeley readers, however.

The two major shortcomings of citation counts for assessing scholarly impact are that they are slow to accumulate and only reflect scholarly impact rather than applied impact. This has led to a need for new metrics to compliment traditional citation metrics (Priem, Piwowar, & Hemminger, 2012). However, many scholars have argued that the new metrics should not be restricted to overcoming the limitations of the previous citation metrics, but can also be expected to provide new insights into research evaluation (Priem & Hemminger, 2010; Torres-Salinas, Cabezas-Clavij & Jemenez-Contrera, 2013).

MENDELEY READERSHIP

Mendeley is an academic social web site for managing references, creating online profiles and sharing with peers. It has an open Applications Programming Interface (API) that can be used for compiling usage indicators with a database of 2.6 million users as of October 2013 (Mohammadi, Thelwall, Haustein, & Larivière, 2015). Currently, Mendeley readership statistics seems to be the most closely related to citation counts, in comparison to other altmetrics. Many studies have used correlations to assess the relationship between Mendeley readership counts and citation counts for the same articles, discovering moderate and positive results. A study of Natural and Science articles published in 2007 shows significant and moderate correlations between Mendeley readership counts and citation counts (Li, Thelwall, & Guistini, 2011). A study of five social science fields with 62,647 articles and five humanities fields with 14,640 found low to moderate significant positive correlations for each discipline (Mohammadi & Thelwall, 2014). This study provides substantial evidence that Mendeley readership could be useful for measuring scholarly impact. Generally, most studies investigating the relationships between Mendeley readers and citation counts (Li & Thelwall, 2012; Haustein, Peters, Bar-Ilan, Priem, Shema, & Terliesner, 2014; Costas, Zahedi & Wouters, 2015) have reported either weak or moderate positive correlations between Mendeley readership counts and citation counts.

Mendeley has a higher proportion of articles with non-zero metrics than other altmetrics (Zahedi, Costas & Wouters, 2014). Out of 19722 publications, 62.6% had at least one reader. Mendeley is particularly used by undergraduates and postgraduates, whereas authors can only make citations. A Mendeley survey found that out of 860 Mendeley users, 55% who had bookmarked articles in Mendeley had actually read them or intended to read and to cite them in their publications (Mohammadi, Thelwall & Kousha, 2015). However, not all readers record their articles in Mendeley, so the data does not actually represent all readers, but most importantly the survey shows that Mendeley bookmark counts seems to be an indicator of readership. An analysis of the 'career stages' of the different Mendeley users found that Postdocs and PhD students register more in Mendeley than any other user category (Zahedi, Costas & Wouters, 2013). Finally, a study of articles that have many or few Mendeley readers compared to their Scopus-indexed citations showed that the reasons can be both technical and theoretical (Thelwall, 2015).

RESEARCH QUESTIONS

The primary goal of this paper is to assess the value of Mendeley readership indicators in conference-based fields. Based on the literature review above, many studies have found significant correlations between Mendeley readers and citation counts and other social media indicators for journal articles. This study focused on four Scopus subject areas: Computer science applications, Computer Software, Building & Construction engineering, and Industrial & Manufacturing engineering, which represent different fields in which conference papers are important. The following research questions drive the study.

1. Do Mendeley readership counts and citation counts reflect the scholarly impact of conference papers Computer Science Applications, Computer Software, Building & Construction Engineering and Industrial & Manufacturing Engineering?
2. Does the answer to the above research question differ between engineering fields and in comparison to journal articles?

METHODS

Correlations between Mendeley readers and citation counts were calculated to ascertain the relationship between readership counts and citation counts for both conference and journal articles. A significant positive correlation gives evidence of a common factor between readership and citation count.

All bibliographic information and citation data for journal articles and conference papers in the four fields from 2011 was extracted from Scopus. The year 2011 was chosen to give a substantial period for citations to accrue, so that there is more chance of getting high correlations between citation and readership counts for both journal articles and conference papers. Spearman correlations were used as the data are skewed. Mendeley reader data was obtained using Webometric Analyst, a free software package. The Mendeley API in Webometric Analyst was used to extract data for Mendeley readers. The Spearman's rank correlation formula was used to calculate 95% confidence intervals for the correlation coefficients. The sampling distribution of the estimate was approximately normal on the transformed scale, hence a 95% CI was found by taking the transformed estimate and adding and subtracting 1.96 times its standard error (Dowdy, Wearden Chilko, 2011. p. 245-246).

RESULTS

Table1 confirms that Mendeley readership counts correlate strongly (0.560-0.662) with citation counts in all subject categories for journal articles. For conference papers, readership counts correlate moderately (0.437-0.439) with citation counts in Computer Science Applications and Software. Readership counts have low correlations (0.143-0.168) with citation counts in Building & Construction and Industrial & Manufacturing Engineering. The reason for the low correlation for conference papers in Building & Construction Engineering and Industrial & Manufacturing Engineering might be due to the low coverage of conference proceedings in engineering subject categories for both Mendeley readers and Scopus citation counts.

Table1. Spearman correlations between Mendeley reader counts and citation counts for articles and conference papers in Scopus from 2011 in the four subject categories analysed. 95% confidence intervals are reported underneath each correlation.

Scopus Subject Category	Articles	Conference papers	Spearman correlation for articles and CI 95%	Spearman correlation for conferences and CI 95%
Computer Science Applications	10000	9999	.560** (0.546,0.573)	.439** (0.423,0.455)
Computer Software	10000	9974	.572** (0.559,0.585)	.437** (0.421,0.453)
Building & Construction engineering	8433	4750	.662** (0.650,0.674)	.143** (0.115,0.171)
Industrial & Manufacturing engineering	10000	9999	.660** (0.649,0.671)	.168** (0.149,0.187)

Table 2 reports the percentage and median level of Scopus citation counts and Mendeley readership counts for both journal articles and conference papers. Computer Science Applications articles mostly (80.9%) have at least one Scopus citation whereas 64.8% have at least one Mendeley reader. In contrast, for conference papers, Computer Software publications attract 68.6% Mendeley readership counts and 54.7% Scopus citation counts, which is the highest percentage of the subject categories analysed.

Table 2. Scopus citation counts and Mendeley readership counts, median and percentage coverage for both journal articles and conference papers.

Scopus subject category	Journal Articles		Conference papers	
	Scopus citations median and % with citations	Scopus Mendeley readership median and % with Mendeley	Scopus citations median and % with citations	Scopus Mendeley readership median and % with Mendeley
Computer Science Applications	3.0 (80.9%)	3.0 (64.8%)	0.0 (34.4%)	0.0 (47.2%)
Computer Software	3.0 (80.6%)	3.0 (62.1%)	3.0 (54.7%)	10.0 (68.6%)
Industrial & Manufacturing Engr.	2.0 (71.3%)	2.0 (55.9%)	0.0 (17.5%)	0.0 (41.0%)
Building & Construction Engr.	2.0 (71.7%)	2.0 (52.7%)	0.0 (18.3%)	-1.0 (18.7%)

DISCUSSION

Mendeley readership counts and Scopus citation counts have strong and significant positive correlations for some fields analysed except, Industrial and Manufacturing Engineering and Building and Construction Engineering, which have weak but positive correlation. Although the findings are restricted to the Scopus citation count, where citations tend to accumulate after several years after the research has been completed and this can only help in long term evaluations.

For journal articles, the strong and positive correlations between Mendeley readership and citation counts for all the studied subject categories of Computer Science Applications, Computer Software, Industrial & Manufacturing Engineering and Building & Construction Engineering corroborate past studies of other areas (Li, Thelwall, & Guistini, 2012; Bar-Ilan, 2012; Mohammadi & Thelwall, 2014).

For conference papers, Computer Software has 68.6% Mendeley readership counts and 54.7% Scopus citation counts. These findings show that the present of conferences is high in Scopus citation counts and Mendeley readership for Computer Science subject category. Whereas, in Building & Construction engineering, conference papers have a much lower percentage coverage; 18.3% and 18.7% for Scopus citation counts and Mendeley readership respectively. This may be due to low coverage of conference proceedings in the field of engineering for Scopus but this cannot explain the results for Mendeley. It may be that a high percentage of engineering conference papers are not of interest to publishing academics, either because of their applied focus or due to disciplinary norms in citation practices.

The percentage and median level of Scopus citation counts and Mendeley readership counts for both journal articles and conference papers are not the same. Computer Science Applications attracted 80.9% of the Scopus citation counts and 64.8% of the Mendeley readership counts in journal articles, which shows high level of citation counts and Mendeley readership counts. While in conference papers, Computer Software attracted 68.6% Mendeley readership counts and 54.7% Scopus citation counts, this show moderate level of coverage. This study shows that Computer Science has wide coverage in both journal articles and conference proceedings.

Based upon high and positive correlations in the subject categories of Computer Science Applications and Computer Software, Mendeley readership counts for conference papers in computer sciences should be acceptable as scholarly impact indicators. In contrast, the weak correlations between Mendeley readership and citation counts in the subject categories of Building & Construction Engineering and Industrial & Manufacturing Engineering, suggest that conference papers in the field of engineering do not support a similar claim.

CONCLUSIONS

This study suggests that Mendeley is rarely used to track academic papers in engineering, but more often in computer science. The high correlation between Mendeley readership and citation counts for conference papers in the field of Computer Science shows that conference proceedings have high enough coverage and counts in Mendeley for readership counts to be a useful impact indicator. In contrast, Mendeley readers provide much weaker evidence of impact in both Building & Construction Engineering and Industrial & Manufacturing Engineering. In conclusion, Mendeley seems to be useful tool for tracking the impact of both conference papers and journal articles in computer science but not for other types of engineering.

REFERENCES

- Bar-Ilan, J. (2012). Jasist 2001–2010. *Bulletin of the American Society for Information Science and Technology*, 38(6), 24-28.
- Bollen, J., Van de Sompel, H., & Rodriguez, M. A. (2008). Towards usage-based impact metrics: First results from the mesur project. Paper presented at the *Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries*, 231-240.
- Bornmann, L., & Marx, W. (2014). The wisdom of citing scientists. *Journal of the Association for Information Science and Technology*, 65(6), 1288-1292.

- Costas, R., Zahedi, Z., & Wouters, P. (2015). Do “altmetrics” correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology*, 66(10), 2003-2019.
- Cronin, B. (2013a). Metrics à la mode. *Journal of the American Society for Information Science and Technology*, 64(6), 1091-1091.
- Dowdy, S., Wearden, S., & Chilko, D. (2011). *Statistics for research*. John Wiley & Sons.
- Drott, M. C. (1995). Reexamining the role of conference papers in scholarly communication. *Journal of the American Society for Information Science*, 46(4), 299.
- Galloway, L. M., Pease, J. L., & Rauh, A. E. (2013). Introduction to altmetrics for science, technology, engineering, and mathematics (STEM) librarians. *Science & Technology Libraries*, 32(4), 335-345.
- Goodrum, A. A., McCain, K. W., Lawrence, S., & Giles, C. L. (2001). Scholarly publishing in the internet age: A citation analysis of computer science literature. *Information Processing & Management*, 37(5), 661-675.
- Gunn, W. (2013). Social signals reflect academic impact: What it means when a scholar adds a paper to mendeley. *Information Standards Quarterly*, 25(2), 33-39.
- Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (2014). Coverage and adoption of altmetrics sources in the bibliometric community. *Scientometrics*, 101(2), 1145-1163.
- Li, X., & Thelwall, M. (2012). F1000, mendeley and traditional bibliometric indicators. Paper presented at the *Proceedings of the 17th International Conference on Science and Technology Indicators*, 2 451-551.
- Li, X., Thelwall, M., & Giustini, D. (2011). Validating online reference managers for scholarly impact measurement. *Scientometrics*, 91(2), 461-471.
- Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010). Altmetrics: A manifesto.
- Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. *ArXiv Preprint arXiv:1203.4745*,
- Mohammadi, E., Thelwall, M., Haustein, S., & Larivière, V. (2015). Who reads research articles? an altmetrics analysis of mendeley user categories. *Journal of the Association for Information Science and Technology*, 66(9), 1832-1846.
- Mohammadi, E., Thelwall, M., & Kousha, K. (2015). Can mendeley bookmarks reflect readership? A survey of user motivations. *Journal of the Association for Information Science and Technology*,

- Mohammadi, E., & Thelwall, M. (2014). Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows. *Journal of the Association for Information Science and Technology*, 65(8), 1627-1638.
- Sud, P. & Thelwall, M. (2014). Evaluating altmetrics. *Scientometrics*, 98(2), 1131-1143. 10.1007/s11192-013-1117-2.
- Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do altmetrics work? twitter and ten other social web services. *PloS One*, 8(5), e64841.
- Thelwall, M. (2015). Why do papers have many mendeley readers but few Scopus-indexed citations and vice versa? *Journal of Librarianship and Information Science*.
- Thelwall, M., & Sud, P. (2015). Mendeley readership counts: An investigation of temporal and disciplinary differences. *Journal of the Association for Information Science and Technology*,
- Torres-Salinas, D., Cabezas-Clavijo, Á., & Jiménez-Contreras, E. (2013). Altmetrics: New indicators for scientific communication in web 2.0. *ArXiv Preprint arXiv:1306.659*.
- Wouters, P., & Costas, R. (2012). *Users, narcissism and control: Tracking the impact of scholarly publications in the 21st century* SURFfoundation Utrecht.
- Zahedi, Z., Costas, R., & Wouters, P. (2013). What is the impact of the publications read by the different mendeley users? Could they help to identify alternative types of impact? Plos alm workshop, San Francisco.
- Zahedi, Z., Costas, R., & Wouters, P. (2014). How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications. *Scientometrics*, 101(2), 1491-1513.



Currencies of Science: discussing disciplinary “exchange rates” for citations and Mendeley readership¹

Rodrigo Costas^{*}, Antonio Perianes-Rodriguez^{**} and Javier Ruiz-Castillo^{***}

[*rcostas@cwts.leidenuniv.nl](mailto:rcostas@cwts.leidenuniv.nl)

Centre for Science and Technology Studies (CWTS), Leiden University P.O. Box 905, Leiden, 2300AX (The Netherlands)

[**antonio.perianes@uc3m.es](mailto:antonio.perianes@uc3m.es)

Departamento de Biblioteconomía y Documentación, SCImago Research Group, Universidad Carlos III, C/ Madrid, 128. 28903 Getafe (Spain)

[***jrc@eco.uc3m.es](mailto:jrc@eco.uc3m.es)

Departamento de Economía, Universidad Carlos III, C/ Madrid, 126. 28903 Getafe (Spain)

INTRODUCTION

In 1998 Garfield stated that “[t]he Mertonian description of normal science describes citations as the currency of science. Scientists make payments, in the form of citations, to their preceptors”. The idea of citations as a currency of science was also discussed by Wouters (1999) who suggested that the “role of the citation might also be compared with that of money, especially if the evaluative use of scientometrics is taken into account. Whenever the value of an article is expressed in its citation frequency, the citation is probably the most important unit of a ‘currency of science’”. Thus, citations have been seen as *currency* able to reward scientists for their work and scientific merit, being an integral part, together with authorship and acknowledgements, of the so-called “reward triangle” (Cronin & Weaver, 1995)².

This role of citations as main currency in evaluative scientometrics has gone unchallenged until recently. The emergence of new ways of measuring the reception of scientific publications by different audiences in the form of the so-called “altmetrics” (Haustein, et al. 2015a; Priem, et al. 2010) probably represents the most important attempt of expanding the system of currencies of science. However, research on altmetrics suggest that there are critical differences with citations: in coverage (Thelwall, et al. 2013), main characteristics (Haustein, et al., 2015), correlations (Costas, et al. 2015b; Haustein, et al. 2014), and interpretation (Haustein et al., 2016). These results essentially highlight the limited potential of most of these metrics as realistic alternatives to citations.

However, *Mendeley readership* has been identified as the most important source or alternative metrics. Mendeley has a high coverage of publications, the correlations between readership and citations are moderate (Zahedi et al., 2014), and readership also has a good filtering ability of highly cited publications (Zahedi, et al. 2015). Consequently, it becomes relevant to

¹ This research project builds on earlier work started by Antonio Perianes-Rodriguez during a research visit to the Centre for Science and Technology Studies (CWTS) of Leiden University as awardee of José Castillejo grant, CAS15/00178, funded by the Spanish MEC. Rodrigo Costas has been partially supported by funding from the DST-NRF Centre of Excellence in Scientometrics and STI Policy (South Africa).

² In an interview in 2012 (Sugimoto, 2012) Blaise Cronin suggested that “[t]he idea of a symbolic capital currency convertor may not be all that far-fetched” regarding the need of determining differential weightings of citations, acknowledgements and other new metrics.

study the characteristics of field Mendeley readership distributions with the same techniques used for studying field citation distributions.

OBJECTIVE

The main objective of this paper is twofold. Firstly, to study whether field Mendeley readership distributions for a large set of Web of Science (WoS) publications are as highly skewed and as similar across fields as it is found for field citation distributions. Secondly, to explore the possibility of estimating “exchange rates” for comparing Mendeley readership counts across fields as has been done for comparing citation counts across them.

METHODOLOGY

Data and analytical approach

We analyse a total of 1,125,811 publications labelled as ‘articles’ in the WoS database, published in 2012, and classified into 30 NOWT disciplines (cf. Ruiz-Castillo & Costas, 2014). Citation scores and Mendeley readership scores³ have been computed.

Two main analytical approaches have been considered:

1) Firstly, we use the Characteristic Scores and Scale technique (CSS hereafter) (Glänzel & Schubert, 1988) to focus on the shape of field citation/readership distributions abstracting from size and scale differences across fields. We partition each citation/readership distribution into three broad categories: (C1) publications with low citations/readership smaller than or equal to μ_l (mean of the metric for the entire distribution); (C2) fairly cited/read publications, with citations/readership greater than μ_l and smaller than or equal to μ_2 (mean of the metric with scores above μ_l); and (C3) publications with remarkable scores greater than μ_2 . Distributions are described by means of two sets of statistics: the percentage of publications in each of the three categories, and the percentages of metrics attributed to the publications in each category. We assess the similarity of the results between fields using the coefficient of variation of the percentages of publications in each category across all disciplines.

2) Secondly, we apply the “exchange rates” approach proposed in Crespo et al (2013)⁴ to study the citation inequality across fields. This approach is based on the comparison of citation counts between fields at a given quantile π . The mean citation of articles belonging to field f and quantile π is denoted by μ_f^π , while the mean citation of articles in that quantile is denoted by μ^π . To express the citations in any field in a given quantile in terms of the citations in a reference situation, we find it useful to define the *exchange rate at quantile π for field f* , $e_f(\pi)$, as follows:

$$e_f(\pi) = \mu_f^\pi / \mu^\pi \quad (1)$$

Note that the citation inequality between the quantities μ_f^π at each quantile, denoted by $I(\pi)$, is entirely attributable to the citation impact differences between the 30 fields holding constant

³ Mendeley REST API has been used for the Mendeley readership data collection (as in Zahedi & van Eck, 2014). Publications with no DOI matching in Mendeley have been considered to have zero readership.

⁴ Crespo et al. (2013) propose that “a field’s citation distribution is like an income distribution in a certain currency”, being able to establish a set of exchange rates that allow the comparison of citations across fields. Here we apply this method to both citations and readership.

the degree of citation excellence in all fields at quantile π . For reasons explained in Crespo et al. (2013), we choose a member of the so-called Generalized Entropy family of inequality indices, which are the only measures of relative inequality that satisfy the usual properties required from any inequality index and, in addition, are decomposable by population subgroup. Similar to previous studies (Crespo, et al. 2013; Albarrán, et al. 2015) the range of the inequality index $I(\pi)$ is particularly high until $\pi \approx 50$, as well as for a few quantiles at the very upper tail of the distributions. However, $I(\pi)$ is essentially constant for a wide range of intermediate values. In this situation, it is reasonable to define an average-based *exchange rate* (ER hereafter) for field f over some quantile interval $[\pi_m, \pi^M]$ in that range as

$$e_f = [1/(\pi^M - \pi_m)] [\sum_{\pi} e_f(\pi)] \quad (2)$$

An advantage of this definition is that we can compute the associated standard deviation (SD), denoted by σ_f . The fact that, for each f , the $e_f(\pi)$ is very similar for all π in the interval $[\pi_m, \pi^M]$ would manifest itself in a small σ_f , and hence in a small coefficient of variation $CV_f = \sigma_f/e_f$.

Distribution of citations and readerships across fields

The results of the CSS analysis for citations and readership are remarkably alike.

Firstly, as illustrated in Figures 1 (citations) and 2 (readership), the distributions are highly skewed in both cases. Specifically, for citations we find that, on average, 69% of all publications belong to C1 and account approximately for 25% of all citations; 22% are in C2, accounting for 34% of all citations; and 9% of the publications are in C3, concentrating 41% of all citations. Virtually the same values are found for readership: 69% of publications in C1 (accounting for 27% of all readership), 22% in C2 (33% of readership) and 9% (41% of readership). These results are comparable with those obtained in previous research concerning citation distributions of scientific fields at different aggregation levels (see *inter alia* Albarrán & Ruiz-Castillo, 2011, Albarrán et al., 2012, Li *et al.*, 2013, and Ruiz-Castillo & Waltman, 2015), research institutions and countries (Glänzel et al. 2014, Perianes-Rodriguez & Ruiz-Castillo, 2015, and Albarrán et al., 2015), and the productivity of authors (Ruiz-Castillo & Costas, 2014). That Mendeley readership is similarly skewed as citations is also supported by Thelwall & Wilson (2015) results for Medical articles.

Secondly, the fairly small coefficients of variation (0.03 for citations and 0.04 for readership in C1, 0.06/0.08 in C2, and 0.13/0.15 in C3) indicate that the distributions of both metrics are extremely similar across fields. This remarkable similarity between fields paves the way for meaningful comparisons of each metric across the 30 scientific fields using the exchange rates defined in equation 2 as normalization factors.

Figure 1. Partition of citation distributions into three categories

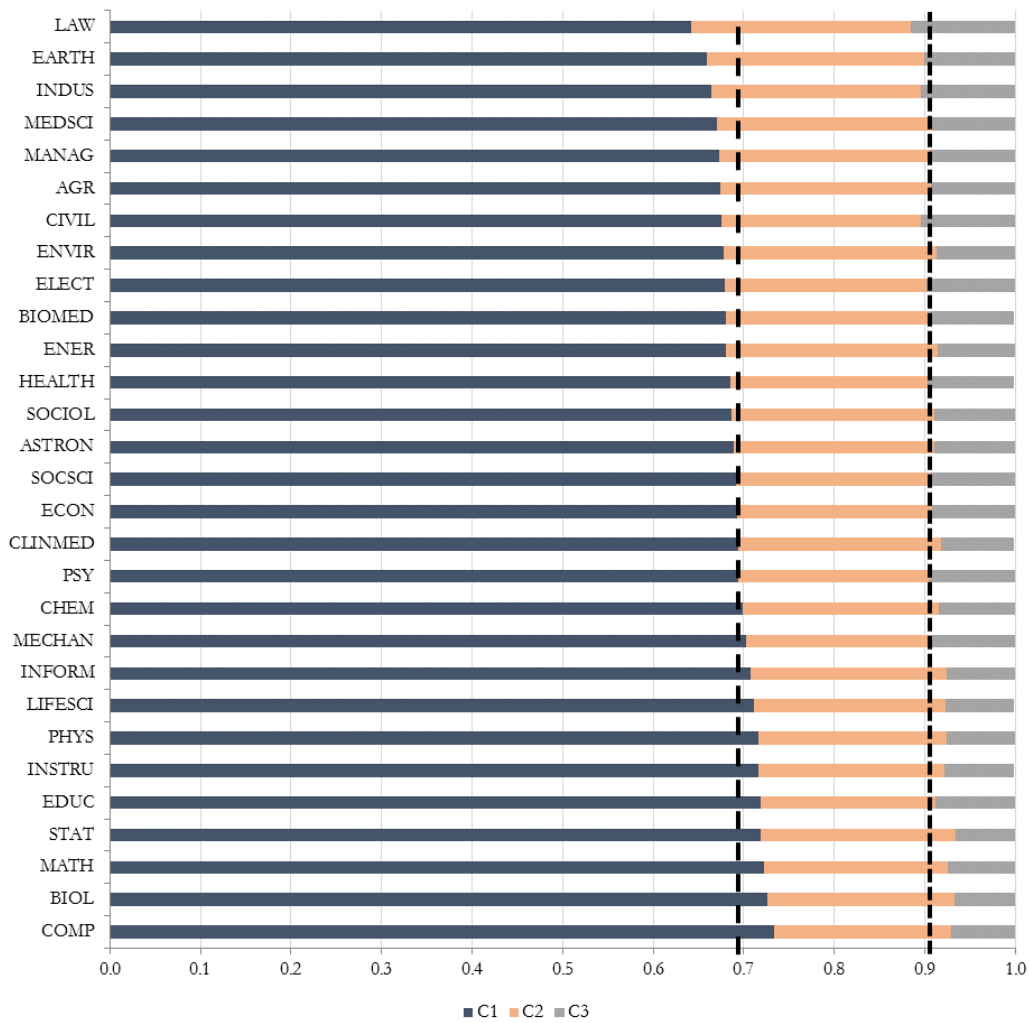
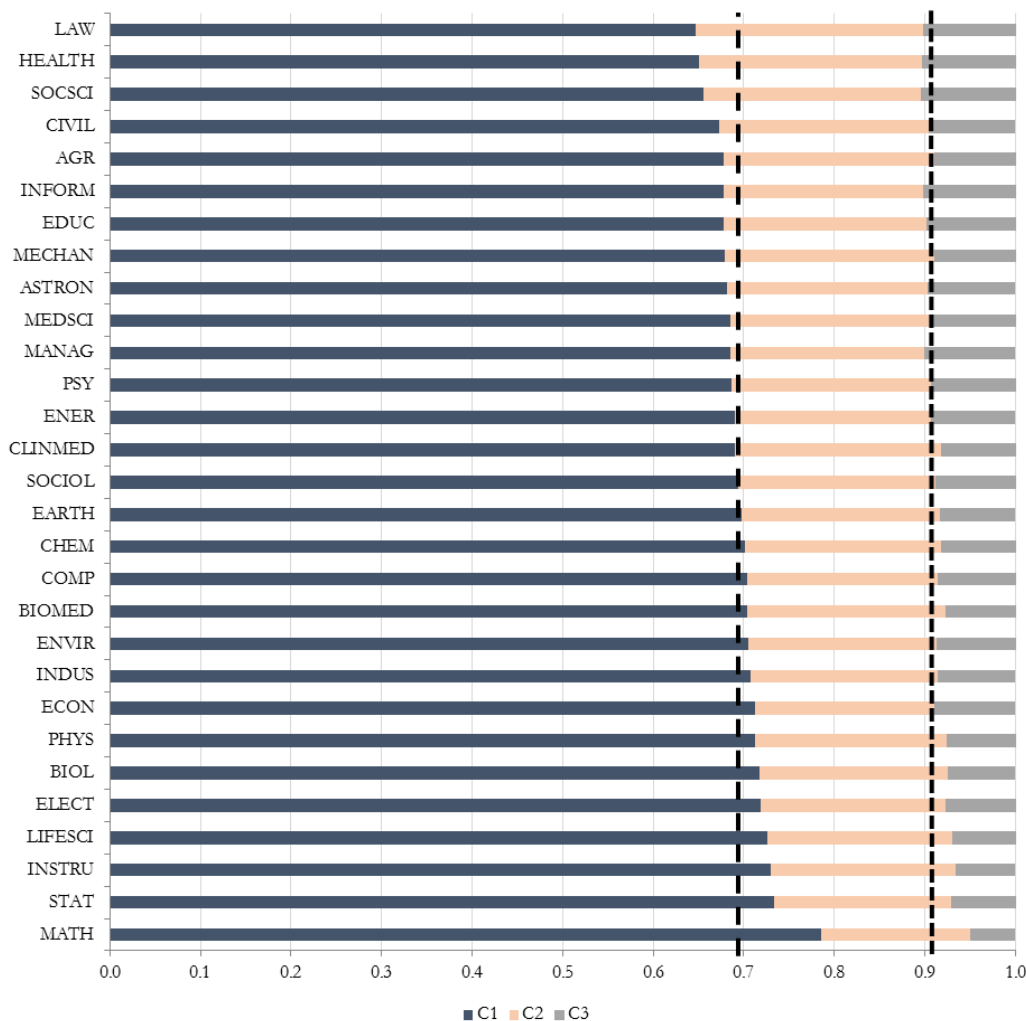


Figure 2. Partition of Mendeley readership distributions into three categories



Comparison of citation and readership across fields

In this study, the inequality index $I(\pi)$ is essentially constant over the percentile interval $[\pi_m, \pi^v] = [50, 97]$, indicating that field citation/readership distributions behave as if they differ by a relatively constant scale factor over a large part of their support. The ERs e_f , as well as the σ_f , and CV_f for citations and readership in that interval are presented in Table 1⁵. On average around 48% of the accumulated of both metrics are included in the interval $[50, 97]$.

We find it useful to divide fields into three groups: Group I (colored in green in Table 1), has a CV_f smaller than or equal to 0.05. This means that the SD of the ER is less than or equal to five percent of the ER itself. Hence, we consider ERs in this group as highly reliable. Group II (yellow), has a CV_f between 0.05 and 0.10. We consider ERs in this group as fairly reliable. Group III (red), with a CV_f greater than 0.10, must be considered as pretty unreliable.

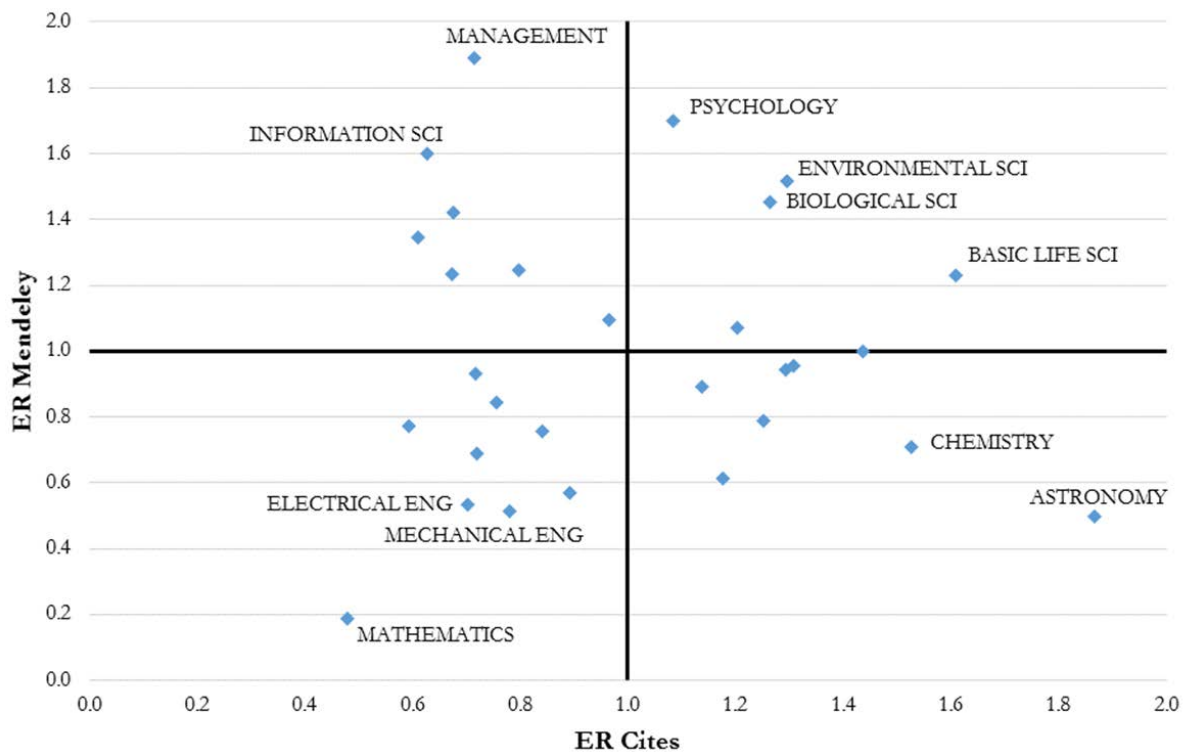
⁵ For example, the first row in Table 1 indicates that 1.14 citations with a standard deviation of 0.08 for an article in *Agriculture and Food Science* between, approximately, the 50th and the 97th percentile of its citation distribution, are equivalent to 1 normalized citation for an article in that percentile interval in the *all-fields case*.

Table 1. Field Exchange Rates over the percentile interval [50, 97]

FIELD	Citations				Mendeley readership			
	ERc	SD	CV	%Total	ERm	SD	CV	%Total
Agriculture and Food Science	1.14	0.08	0.07	54.57	0.89	0.03	0.04	42.76
Astronomy and Astrophysics	1.87	0.04	0.02	89.57	0.50	0.02	0.04	23.98
Basic Life Sciences	1.61	0.07	0.04	77.24	1.23	0.06	0.05	58.98
Basic Medical Sciences	1.31	0.08	0.06	62.82	0.96	0.02	0.02	45.95
Biological Sciences	1.26	0.04	0.03	60.70	1.45	0.03	0.02	69.79
Biomedical Sciences	1.44	0.06	0.05	68.91	1.00	0.02	0.02	47.96
Chemistry and Chemical Engineering	1.53	0.03	0.02	73.23	0.71	0.03	0.05	34.00
Civil Engineering and Construction	0.84	0.04	0.05	40.39	0.76	0.02	0.02	36.32
Clinical Medicine	1.25	0.03	0.03	60.12	0.79	0.03	0.04	37.81
Computer Sciences	0.72	0.06	0.08	34.45	0.93	0.07	0.08	44.68
Earth Sciences and Technology	1.20	0.05	0.04	57.82	1.07	0.02	0.02	51.45
Economics and Business	0.68	0.06	0.08	32.41	1.42	0.08	0.06	68.20
Educational Sciences	0.61	0.05	0.08	29.27	1.35	0.08	0.06	64.62
Electrical Engineering and Telecommunication	0.70	0.06	0.08	33.73	0.53	0.06	0.11	25.56
Energy Science and Technology	1.29	0.03	0.02	62.04	0.94	0.03	0.03	45.29
Environmental Sciences and Technology	1.29	0.04	0.03	62.16	1.51	0.04	0.02	72.69
General and Industrial Engineering	0.72	0.04	0.05	34.55	0.69	0.04	0.06	33.15
Health Sciences	0.96	0.04	0.05	46.30	1.09	0.11	0.10	52.48
Information and Communication Sciences	0.63	0.05	0.07	30.08	1.60	0.06	0.03	76.83
Instruments and Instrumentation	0.89	0.05	0.06	42.89	0.57	0.03	0.06	27.36
Law and Criminology	0.59	0.04	0.07	28.48	0.77	0.08	0.10	37.08
Management and Planning	0.71	0.04	0.06	34.31	1.89	0.06	0.03	90.71
Mathematics	0.48	0.05	0.10	23.06	0.19	0.05	0.28	9.09
Mechanical Engineering and Aerospace	0.78	0.04	0.05	37.49	0.51	0.02	0.04	24.72
Physics and Materials Science	1.18	0.07	0.06	56.43	0.61	0.06	0.09	29.47
Psychology	1.08	0.04	0.04	52.07	1.70	0.12	0.07	81.62
Social and Behavioral Sciences, Interdisciplinary	0.80	0.04	0.04	38.30	1.25	0.11	0.09	59.79
Sociology and Anthropology	0.67	0.05	0.07	32.33	1.23	0.06	0.05	59.19
Statistical Sciences	0.76	0.05	0.06	36.32	0.84	0.11	0.13	40.45

The ERs for most fields are reliable or fairly reliable in terms of their coefficient of variation, particularly regarding citations. The less reliable case is *Mathematics* both in terms of citations and Mendeley readership.

There are fields with high ERs for citations and relatively low for readership and vice versa (see Figure 3), this being obviously related to the higher (lower) densities of citation or readership that can be found across fields. Social Sciences fields such as *Management and Planning* have high ER for readership and lower for citations. Fields like *Astronomy & Astrophysics* represent the contrary pattern, higher ERs for citations and lower for readership. *Mathematics* is an example of low ERs in both metrics, while *Basic Life Sciences* is an example of high ERs in both metrics. These results align with previous results (Costas et al., 2015a; Haunschild & Bornmann, 2016) on the existence of different rankings of disciplines based on the abundance/scarcity of one metric or the other.

Figure 3. Correlation ER Citations vs. ER Mendeley readership

Crespo et al (2013, p. 6) suggested that in order to cardinaly compare the citation impact of two publications, the raw citations could be normalized against their ERs. In our case, based on the ERs reported in Table 1, for example a paper in *Educational Sciences* with 2 citations would be normalized to $2/0.61 = 3.28$, while an *Astronomy and Astrophysics* paper with 3 citations would be normalized to $3/1.87=1.60$. By dividing the two normalized values ($3.28/1.60 = 2$) we conclude that the *Educational Sciences* paper has an impact 2 times higher than the *Astronomy and Astrophysics* paper. Focusing on readership the two fields show an inverse pattern. In *Educational Sciences* 1.35 Mendeley readership would be exchanged by one item of normalized readership impact, while for *Astronomy & Astrophysics* it would require only 0.50 readership. The obvious reason is that Mendeley readership has a higher density in *Educational Sciences* than in *Astronomy and Astrophysics*, so that high values of readership in *Educational Sciences* will be equivalent to relatively lower levels of readership in *Astronomy and Astrophysics*.

A potential question that may arise is whether it is valid to compare the normalized values of citations and readership based on their ER. To answer this question let's take the example of *Astronomy & Astrophysics*. By naïvely applying the current ERs one could conclude that, given the scarcity of readership in this field, one Mendeley readership would be worth more than one citation in terms of normalized impact⁶. However, such interpretation obviously does not take into account the intrinsic *value* of the two 'currencies' in terms of measuring 'symbolic capital' (Bourdieu, 1986). Probably, given the choice, most scientists would prefer to receive citations instead of (only) readership. A plausible explanation for this higher esteem for citations over readership can be found in the framework of acts around research objects

⁶ The ERs would indeed suggest that 1 citation would only provide $1/1.87 = 0.53$ of normalized citation impact, while 1 readership would provide $1/0.50 = 2$ of normalized readership impact.

proposed by Haustein et al (2016). There, readership appear as acts of “access”, related to a lower level of “engagement” by the users; while citations are seen as acts of “appraisal” related to a stronger engagement with the cited publication⁷. Thus, taking engagement as a crude proxy of the potential value attached to the metrics by different actors, it can be argued that these two metrics can be easily perceived to have different values to measure symbolic capital⁸, therefore making their direct comparison (even with normalized scores) incommensurable.

CONCLUSIONS

The most important conclusion of this paper is that, given the strikingly similarity of Mendeley readership distributions between fields, it is technically feasible and reasonable to estimate field ERs for readership over a wide range of intermediate quantiles, in a similar fashion as it has been done for citations. Such ERs could be used for comparing readership counts between fields, as well as reducing the effect on overall readership inequality of the idiosyncratic readership differences between fields⁹.

Our results also raise the important question related to the different abilities of different metrics in capturing scientific impact or symbolic capital. Further research should focus on discussing the different “values” of metrics (i.e. citations and readership, but also other social media metrics) for measuring research performance, thus contributing to determine their potential complementarities and exchangeabilities and allowing their proper consideration as new “currencies” of science.

References

- Albarrán, P. and J. Ruiz-Castillo (2011). References Made and Citations Received By Scientific Articles. *Journal of the American Society for Information Science and Technology*, 62, 40-49.
- Albarrán, P., J. Crespo, I. Ortuño, and J. Ruiz-Castillo (2011). The Skewness of Science In 219 Sub-fields and a Number of Aggregates. *Scientometrics*, 88, 385-397.
- Albarrán, P., Perianes-Rodríguez, A., & Ruiz-Castillo, J. (2015). Differences in Citation Impact Across Countries. *Journal of the Association for Information Science and Technology*, 66(3), 512–525. doi:10.1002/asi
- Bourdieu, P. (1986). The forms of capital. In *Handbook of Theory and Research for the Sociology of Education* (Vol. 241, pp. 241–258). doi:10.1002/9780470755679.ch15
- Costas, R., Zahedi, Z., & Wouters, P. (2015a). The thematic orientation of publications mentioned on social media: large-scale disciplinary comparison of social media metrics

⁷ Of course there can be deviations to this pattern, with users having a stronger engagement in Mendeley than with citations (e.g. some papers may be thoroughly read in Mendeley while some others may be only superficially cited).

⁸ The different perceived values and esteem of new social media metrics in terms of symbolic capital are illustrated in Desrochers et al. (2016).

⁹ Field-normalized indicators based on Mendeley (e.g. Mean Normalized Readership Scores) have been already proposed (Haunschild & Bornmann, 2016).

- with citations. *Aslib Journal of Information Management*, 67(3), 260–288.
doi:10.1108/AJIM-12-2014-0173
- Costas, R., Zahedi, Z., & Wouters, P. (2015b). Do “altmetrics” correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology*, 66(10), 2003–2019.
- Crespo, J. a, Li, Y., Li, Y., & Ruiz-Castillo, J. (2013). The measurement of the effect on citation inequality of differences in citation practices across scientific fields. *PLoS one*, 8(3), e58727. doi:10.1371/journal.pone.0058727
- Cronin, B., & Weaver, S. (1995). The praxis of acknowledgement: from bibliometrics to influmetrics. *Revista Española de Documentación Científica*, 18(2), 172–177.
- Desrochers, N., Paul-Hus, A., Haustein, S., Costas, R., Mongeon, P., Quan-Haase, A., ... Larivière, V. (2016). Authorship, inventorship, citations, acknowledgements, and social media: symbolic capital in the multifaceted Reward System of Science. *Under review*.
- Garfield, E. (1998). The Use of Journal Impact Factors and Citation Analysis. In *Cell Separation, Hematology and Journal Citation Analysis Mini Symposium in tribute to Arne Bøyum* (pp. 1–15). Retrieved from http://www.garfield.library.upenn.edu/papers/eval_of_science_oslo.html
- Glänzel, W.; Schubert, A. (1988). Characteristics scores and scales in assessing citation impact. *Journal of Information Science*, 14(2), 123-127
- Glänzel, W. Thijs, B., and Debackere, K. (2014). The application of citation-based performance classes to the disciplinary and multidisciplinary assessment in national comparison and international research assessment. *Scientometrics*, 101 , 939-952.
- Haunschild, R., & Bornmann, L. (2016). Normalization of Mendeley reader counts for impact assessment. *Journal of Informetrics*, 10(1), 62–73. doi:10.1016/j.joi.2015.11.003
- Haustein, S., Bowman, T. D., & Costas, R. (2016). Interpreting “altmetrics”: viewing acts on social media through the lens of citation and social theories. In C. R. Sugimoto (Ed.), *Theories of Informetrics: A Festschrift in Honor of Blaise Cronin* (pp. 372–405). Berlin: De Gruyter Mouton. Retrieved from <http://arxiv.org/abs/1502.05701>
- Haustein, S., Costas, R., & Larivière, V. (2015). Characterizing social media metrics of scholarly papers: the effect of document properties and collaboration patterns. *PLoS ONE*. doi:10.1371/journal.pone.0120495
- Haustein, S., Peters, I., Sugimoto, C. R., Thelwall, M., & Larivière, V. (2014). Tweeting Biomedicine □: An Analysis of Twitter. *Journal of the Association for Information Sciences and Technology*, 65(4), 656–669.
- Li, Y., Radicchi, F., Castellano, C., & Ruiz-Castillo, J. (2013). Quantitative Evaluation of Alternative Field Normalization Procedures. *Journal of Informetrics*, 7, 746–755.

- Perianes-Rodríguez, A., and Ruiz-Castillo, J. (2016). University citation distributions. Forthcoming in *Journal of the American Society for Information Science and Technology* (DOI: 10.1002/asi.23619)
- Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010). Alt-metrics: a manifesto. <http://altmetrics.org/manifesto/>.
- Ruiz-Castillo, J., & Costas, R. (2014). The skewness of scientific productivity. *Journal of Informetrics*, 8(4), 917–934. doi:10.1016/j.joi.2014.09.006
- Ruiz-Castillo, J., & Waltman, L. (2015). Field-normalized citation impact indicators using algorithmically constructed classification systems of science. *Journal of Informetrics*, 9, 102-117.
- Sugimoto, C. R. (2012). Taking the Measure of Metrics: Interviews with Four ASIS&T Members. *Bulletin of the American Society for Information Science and Technology*, 38, 33–38.
- Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do Altmetrics Work? Twitter and Ten Other Social Web Services. *PLoS ONE*, 8(5), e64841. doi:10.1371/journal.pone.0064841
- Thelwall, M., & Wilson, P. (2015). Mendeley Readership Altmetrics for Medical Articles: An Analysis of 45 Fields. *Journal of the Association for Information Science and Technology*. doi:10.1002/asi.23501
- Wouters, P. (1999). *The Citation Culture*. University of Amsterdam.
- Zahedi, Z., Costas, R., & Wouters, P. (2014). How well developed are altmetrics? A cross-disciplinary analysis of the presence of “alternative metrics” in scientific publications. *Scientometrics*, 101, 1491–1513. doi:10.1007/s11192-014-1264-0
- Zahedi, Z., Costas, R., & Wouters, P. (2015). Do Mendeley readership counts help to filter highly cited WoS publications better than average citation impact of journals (JCS)? *arXiv:1507.02093*. Retrieved from <http://arxiv.org/abs/1507.02093>
- Zahedi, Z., & van Eck, N. J. (2014). Visualizing readership activity of Mendeley users using VOSviewer. In *altmetrics14: expanding impacts and metrics*. Bloomington, Indiana, USA: An ACM Web Science Conference 2014 Altmetrics Workshop. doi:10.6084/m9.figshare.1041819



SSH & the City. A Network Approach for Tracing the Societal Contribution of the Social Sciences and Humanities for Local Development ¹

Nicolas Robinson-Garcia^{*}, Thed N. van Leeuwen^{**} and Ismael Rafols^{***}

^{*}elrobin@ingenio.upv.es

INGENIO (CSIC-UPV), Universitat Politècnica de València, Valencia (Spain)

^{**}leeuwen@cwts.leidenuniv.nl

Centre for Science and Technology Studies (CWTS), Leiden University, Leiden (The Netherlands)

^{**}i.rafols@ingenio.upv.es

INGENIO (CSIC-UPV), Universitat Politècnica de València, Valencia (Spain)

SPRU, University of Sussex, Brighton (United Kingdom)

Observatoire des Sciences et Techniques, HCERES, Paris (France)

ABSTRACT

Current evaluation frameworks in research policy were designed to address: 1) life and natural sciences, 2) global research communities, and; 3) scientific impact. This is problematic, as they do not adapt well to SSH scholarship, to local interests, or to consider broader societal impacts. This paper discusses three different evaluation frameworks and proposes a methodology to operationalize them and capture societal interactions between social sciences and humanities (SSH) researchers and their local context. To capture such interactions, we propose the use of social media and web-link analysis to identify interactions between academics and local stakeholders. We consider that the power of these tools is not so much on understanding their meaning as ‘acts’ to develop impact or visibility metrics whenever a mention to a research article is made, but as proxies for personal interactions. We offer some examples of the expected social networks we aim at developing for two Spanish cities: Granada and Valencia.

INTRODUCTION

Current evaluation frameworks in research policy were designed to address: 1) life and natural sciences, 2) global research communities, and; 3) scientific impact. This is problematic, as they do not adapt well to SSH scholarship, to local interests, or to consider broader societal impacts. Moreover, their focus on outputs, implicitly assumes a linear relation between the activity (research) and the expected result (publication). This linear perspective is particularly problematic to the SSH areas as they can address different audiences (Nederhof 2006). But many of the ‘impacts’ SSH activity may have in society are due to multiple factors (problem of attribution) and are of a secondary nature that is, the outcome may not be traceable to any single given output but to cascading effects (Upton et al. 2014). Here we propose a network approach for identifying societal contributions in local contexts. The goal of this approach is not to develop indicators for benchmarking, but to map interactions for strategic assessment.

¹ Nicolas Robinson-Garcia is supported by a Juan de la Cierva-Formación Fellowship granted by the Spanish Ministry of Economy and Competitiveness.

The absolute “value” or “weight” of the interactions cannot be captured, but we hope the method can identify the hot spots where they are taking place.

This paper discusses three different evaluation frameworks and proposes a methodology to operationalize them and capture societal interactions between social sciences and humanities (SSH) researchers and their local context. Many countries are putting an increasing emphasis on the societal impact discourse when assessing research performance. They are calling for evidences of societal impact, urging researchers to engage on social outreach and public engagement. Citizen science (Irwin 1995), science for the people (Layton 1973), mode 2 science and society (Nowotny et al. 2001), public value of science (Brewer 2013) or societal impact (Spaapen & Drooge 2011) are just some examples of the concepts used indistinctively to relate to the co-evolving interactions that develop between society and science.

CONCEPTUAL FRAMEWORKS

To avoid linear “impact” indicators, we suggest the application of alternative evaluation frameworks which consider process-based indicators. Because output-based approaches consider the relation between the activity and the outcome to be linear, it ties researchers to take an expected course of action, intruding in many cases in their communication patterns (Fuchs 2014). A process-based approach encourages researchers towards social engagement without tying them to a closed set of outputs. This means a shift from an ‘impact’ discourse to promoting ‘productive interactions’ between academics and non-academics (Molas-Gallart & Tang 2011).

Here we consider three frameworks which, to our understanding, define and describe in a consistent manner process-based interactions, but have so far failed at proposing a scalable and quantitative methodology to assess them. These frameworks intend to overcome problems raised when trying to assess scholarly performance and activities beyond scientific impact. The first approach we refer to is the ‘third stream metrics’ (Molas-Gallart et al. 2002) where they differentiate between activities and capabilities of universities. Capabilities are defined as of two types: physical facilities and knowledge capabilities. Based on this key differentiation a set of indicators is developed associated to each aspect. The second framework is based on the ‘productive interactions’ concept (Spaapen & van Drooge 2011), which adopts a process-based perspective. Its originality is due to explicitly driving away from an ‘impact’ discourse (understood as the effect of research in society).

If we look at the ‘third stream metrics’ framework and the ‘productive interactions’ concept in terms of networks, we could make the following analogy. Capabilities represent nodes; and activities, defined as productive interactions, are the edges of a network by which academics and institutions interact and exchange knowledge with non-academics. Here is where the third framework, the ‘knowledge value’ framework (Rogers & Bozeman 2001), becomes relevant. This framework aims at defining the unit of analysis that should be used for evaluation instead of R&D programs. They present two core concepts: knowledge value collectives (KVC) and knowledge value alliances (KVA). In short, KVC is defined as the set of individuals who share a common knowledge base, while KVA is defined as a subset of a KVC where individuals are interacting with each other.

EXPECTED METHODOLOGICAL CONTRIBUTION

In our understanding, the three proposals referred to above describe the research process and collectives as interactions, flows and connections, but fail to recognize the network approach as a way of implementing them empirically. Instead, they rely on unidimensional indicators. Indeed, it was Harrison White the one to first relate peoples' activities with their interactions and the importance of context on the shaping of such interactions (Watts 2004). Visualizing societal contributions as processes instead of societal impact as outcomes presents important advantages for research evaluation practices. It allows policy makers to take strategic action and to try to anticipate the desired impact (i.e., offering 'institutional' support or promoting universities' role in society). Also networks of people and institutions can be seen as forms of embodied cultural and social capital. Due to the richness of the connections and the heterogeneity of its different actors, they have a value by themselves which should be considered when assessing the role of universities in local development.

To capture such interactions, we propose the use of social media and web-link analysis to identify interactions between academics and local stakeholders. We consider that the power of these tools is not so much on understanding their meaning as 'acts' to develop impact or visibility metrics whenever a mention to a research article is made (Haustein et al. 2015), but as *proxies* for personal interactions. These networks are what Rogers & Bozeman (2001) refer to as KVA. We hypothesize that such alliances may or may not be established within a KVC. We argue that both types of alliances may be established. This methodology may benefit SSH areas, due to the direct and informal nature of relations between scholars from these fields and non-academics (Olmos-Peñuela et al. 2014). The strength of such approach is based on its potential to monitor unstructured interactions, as well as those which are the result of a specific research action (i.e., an R&D project). Therefore, research policy makers can monitor and better comprehend the process of interactions developed, and also identify other hot spots where productive interactions may well be taking place.

PRELIMINARY RESULTS AND FURTHER STEPS

In order to test our methodology, we will conduct several case studies in different cities of Spain and The Netherlands focusing for each of them on specific events and SSH areas. Here, we show two examples of the types of networks we aim at discovering through our approach. These examples are very preliminary results and should be considered as illustrations of the methodology rather than complete analyses of our case studies.

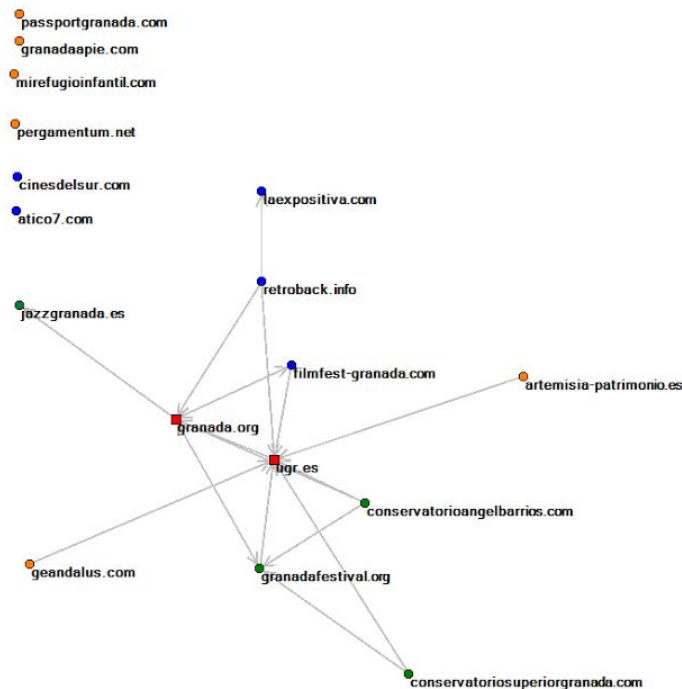
Figure 1 shows some preliminary results of relations between a sample of spin-offs, music-related and movie-related institutions in the city of Granada. As observed, although interactions through the web are not fully explored in this first approach, we can already capture ties between cultural events and institutions and the university. We also observe the strength of the tie between the city and the university, highlighting its role as an anchor institution for cultural life of the city (Goddard et al. 2014).

In our second test (figure 2) we focus in the case of the city of Valencia and web-links between university and a sample of local associations and cultural events. A complete different picture emerges here. First, we find three components instead of one: one formed by a neighbourhood association and a local association, a second one where the University of Valencia (uv.es) acts as central node connecting a local political association and other associations defending local agriculture, and a third component relating theatre and arts

institutions and associations. Second, we observe how three of the four universities analysed (ucv.es, uchceu.es and upv.es) remain as isolated nodes of the network. In this case, the University of Valencia seems to be the only one establishing ties with local bodies, although its role in our network does not seem to be as central as in the former case. Interestingly, local authorities such as the local council (valencia.es) the regional department of culture (culturartsgeneralitat.es) do not play a crucial role either.

Although these examples need further refinement, they offer a good example of the type of interactions we are proposing to capture through our network approach. In this case we have used an out-link analysis between a selected sample of institutional websites without going into much depth in our analyses. Future directions will go into the designing and analysis on the relation between the university and specific events or institutions in the city. Our aim is to go beyond an institutional perspective and make use of Twitter, blogs and Facebook to identify direct personal interactions that may reflect a greater (but also informal) role of universities in the case of Valencia for instance, than previously noted through our web-link analysis.

Figure 1. Out-link network of the University of Granada and its interactions with other local institutions. Node colours: red > University of Granada and City Council; green > music related festivals and institutions; orange > spin-offs; blue > movie related festivals and institutions. Depth of crawling: 1.

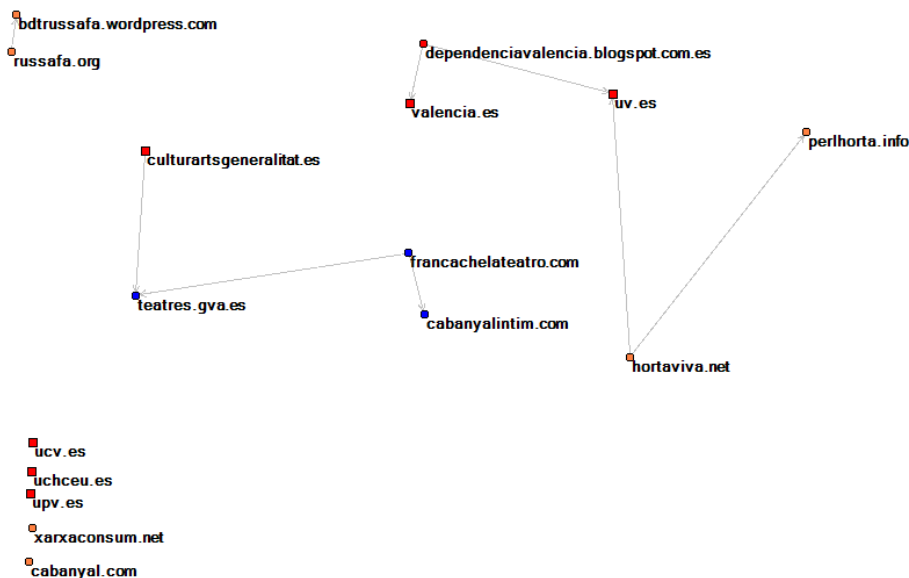


In the case of Twitter data, we expect to encounter a series of challenges derived from the nature of this social media platform. These are the following:

1. Twitter data is retrieved by querying its API, which means that our network will be dependent on such query and we will not know to what extent we are showing a complete picture of the activity we are querying.

2. Because of such difficulty to retrieve and manage complete datasets, we cannot analyse local interactions between academics and non-academics in general, but must focus on specific events. Here we define events as the sociocultural activity or movement of our interest.
3. Due to the size of the network and the informal nature of the platform, nodes representing academics are not self-evident or easily identifiable.
4. It is equally difficult to identify non-academics having an active role in a given activity.
5. Even after having identified non academics and academics, we should be able to establish through the network the level of engagement of these two groups in a given discussion (in our case, related with a given sociocultural local event).

Figure 2. Out-link network of four Valencian universities and their interactions with other local institutions. Node colours: red > universities and local and regional public institutions; green > music related festivals and institutions; orange > social and neighbourhood associations; blue> theatre and arts related associations. Depth of crawling: 1.



However, Twitter has the positive aspect of informing us of different types of ties between nodes, hence links can be established in terms of followers and followees, mentions or retweets. Such distinctions allow us to distinguish between social distance and network paths (Watts, Dodds & Newman, 2002). That is, two dimensions of the network which could allow us to identify 1) potential academics, and 2) potential actors involved in a given event. We define social distance as that related to the acknowledgement two nodes make of each other. Hence, if node A and node B follow each other, we consider them ‘socially’ close to each other. Network paths are defined as those which link two nodes by discussing common topics.

Academic nodes are identified in terms of social distance. This can be done following two possible strategies:

- 1) We query for a given university's Twitter account and identify all its followers and followees. We consider academics all those nodes which follow the university and are followed by the university.
- 2) We identify through Altmetric.com API all Twitter accounts discussing research papers authored by a given university. We then cross these accounts with the university account and consider those nodes to be local and academic. This definition would be more restrictive.

In order to identify network paths between academics and non-academics with regard to a given local sociocultural event, we query the Twitter API in order to retrieve discussions related to such event. Then we locate previously identified academic nodes and analyse through network centrality measures their role in the discussion and how they relate with non-academic nodes.

CONCLUDING REMARKS

The present paper proposes a novel approach for analysing societal contributions of SSH academics to their cities. Although the methodology is not fully implemented, very preliminary results are offered for illustrating its potentialities. This proposal shows a quantitative approach which can be of use for research policy makers. Such methodology is characterized by the following aspects:

1. *Local versus global impact.* It is directed at the contribution academics make to local development, as opposed to traditional indicators such as citation metrics.
2. *Sociocultural impact of academia.* Societal relevance is traditionally considered in socioeconomic terms. Focusing on interactions rather than impact-based indicators offers a wider perspective as to what is considered 'societal'.
3. *Social networks instead of impact indicators.* As illustrated by the three evaluation frameworks presented above, recent developments in research evaluation are directed at analysing interactions between institutions, academics and other actors (firms, non-academics, society in general). However, they fail at operationalising their proposals in terms of network analysis. Network analysis and mapping have already been proven a powerful tool for research evaluation (Noyons, 2005; Wallace & Rafols, 2015), however they have not yet been applied in the context of societal impact.
4. *Social media as a data source.* Although other studies have analysed societal contributions in SSH, they usually recollect their data either through surveys (Olmos-Peñuela, Castro-Martínez & D'Este, 2014) or interviews (Olmos-Peñuela et al., 2014). Such methods are of limited success for research policy purposes for the following reasons: 1) they are costly in terms of time and money, 2) they are highly dependent on the subject's capacity to inform fully of their activity, and 3) they are intrusive, intervening with academics' activities. Using web-link analysis and, specially, Twitter, offers a relatively easy, non-intrusive and 'decontextualized' way of retrieving data for analysing informal interactions.

REFERENCES

- Brewer, J.D. (2013). *The Public Value of the Social Sciences: An Interpretive Essay*. A&C Black.
- Fuchs, M.Ž. (2014). Bibliometrics: Use and abuse in the humanities. Blockmans W., Engwall, L., & Weaire D. (eds) *Bibliometrics. Use and Abuse in the Review of Research Performance*, pp. 107–16. Portland Press: London.
- Goddard, J., Coombes, M., Kempton, L., & Vallance, P. (2014). Universities as anchor institutions in cities in a turbulent funding environment: vulnerable institutions and vulnerable places in England. *Cambridge Journal of Regions, Economy and Society*, 7(2), 307–25.
- Haustein, S., Bowman, T.D., & Costas, R. (2015). Interpreting ‘altmetrics’: viewing acts on social media through the lens of citation and social theories. *arXiv:1502.05701*.
- Irwin, A. (1995). *Citizen Science. A Study of People, Expertise and Sustainable Development*. London and New York: Routledge.
- Layton, D. (1973). *Science for the People*. London: Allen & Unwin.
- Molas-Gallart, J., Salter, A., Patel, P., Scott, A., & Duran, X. (2002). Measuring third stream activities. *Final report to the Russell Group of Universities*. Brighton: SPRU, University of Sussex.
- Molas-Gallart, J., & Tang, P. (2011). Tracing ‘productive interactions’ to identify social impacts: an example from the social sciences. *Research Evaluation*, 20(3), 219–26.
- Nederhof, A.J. (2006). Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A Review. *Scientometrics*, 66(1), 81–100.
- Nowotny, H., Scott, P., & Gibbons, M. (2001). *Re-thinking science: knowledge and the public in an age of uncertainty*. Cambridge: Polity Press.
- Olmos-Peñuela, J., Molas-Gallart, J., & Castro-Martínez, E. (2014). Informal collaborations between social sciences and humanities researchers and non-academic partners. *Science and Public Policy*, 41(4), 493–506.
- Rogers, J.D., & Bozeman, B. (2001). ‘Knowledge Value Alliances’: An Alternative to the R&D Project Focus in Evaluation. *Science, Technology, & Human Values*, 26(1), 23–55.
- Spaapen, J., & van Drooge, L. (2011). Introducing ‘productive interactions’ in social impact assessment. *Research Evaluation*, 20(3), 211–8.
- Upton, S., Vallance, P., & Goddard, J. (2014). From outcomes to process: evidence for a new approach to research impact assessment. *Research Evaluation*, 23(4), 352–65.
- Watts, D.J. (2004). *Six degrees: The science of a connected age*. Vintage.
- Watts, D.J., Dodds, P.S., & Newman, M.E.J. (2002). Identity and search in social networks. *Science*, 296(5571), 1302-1305.



Comparing the characteristics of highly cited titles and highly alted titles

Fereshteh Didegah¹, Timothy D. Bowman¹, Sarah Bowman¹ & James Hartley²

¹Research Unit for the Sociology of Education
University of Turku, Finland

²School of Psychology
Keele University, Staffordshire, UK

ABSTRACT

This study examines differences in the types of titles for articles that show high altmetric activity (highly alted articles) versus highly cited articles. This work expands on previous research on document titles in combination with a grounded theory approach to develop a codebook in which articles were manually coded based on 11 characteristics. The results show that there are differences and similarities in titles across many of the examined characteristics; highly cited titles and highly mentioned titles on Wikipedia have some similar characteristics such as they have the the highest percentage of substantive words; in addition, there are no or very few titles referencing outside or with humor/lightness on both platforms. Twitter and Facebook also showed some similarities having the highest percentage of humorous/light titles and lowest percentage of substantive words in their titles.

INTRODUCTION

Titles have been shown to be a very important component of a document. A document title is one of the first points of contact between an article and its readers, as most readers decide whether or not to read the entire article after reading its title. The ways in which authors compose titles can bring attention to a document through the use of different strategies including humor, title length, punctuation, or idioms (Subotic & Mukherjee, 2014). Receiving attention for scientific work, such as readership and citations, is highly competitive and considered an essential aspect of the academic reward system (Cronin, 1996).

The consumption and dissemination of science is taking place both online and offline. Traditional bibliometric analyses techniques are used when analyzing citations (e.g., Merton, 1973; Cronin, 1984), whereas digital traces of scientific document mentioned in online contexts such as Twitter or Facebook, news sites, blogs, and other platforms are examined by webometric (Björneborn & Ingwersen, 2004) and (more recently) altmetric researchers.

Within the bibliometrics community, there have been many papers examining document title types and citation activity. Subotic and Mukherjee (2014) examined the impact of title amusement on the number of article downloads and found that titles with a high level of amusement have a higher number of downloads, but that this did not correlate with the number of citations to the articles. Articles that had titles that posed questions were also downloaded at a high rate, but cited less. In another work, descriptive titles were found to be downloaded and also cited more than articles with declarative titles (Jamali & Nikzad, 2011).

When examining document titles in the field of psychology, authors (Haslam et al., 2008) found that papers with longer titles received fewer citations than others. In contrast, it was found that title length, as measured by the number of significant words in the title, did not associate with citations to articles in marketing (Stremersch, Verniers and Verhoef, 2007). In earlier work examining law reviews, it was found that shorter titles received a higher number of citations (Ayres & Vars, 1999). Interestingly there have been no macro-studies that have considered title length simultaneously with other factors, such as abstract length. It may be that longer titles cannot easily draw a reader's attention to the main message of the article, whereas a shorter title could speed decisions about article relevancy.

These numerous bibliometric results suggest that title types are important and worthy of study in the domain of altmetrics. As the attention of scholars is being consumed by the vast amount of articles being published and because garnering citations and readership are critical components of the academic reward system, it is important to consider what types of titles collect the most amount of attention in the online world. This study looks to shed light on the effect of title types on Twitter, Weibo, Facebook, Wikipedia, blogs, and news activity. The title types on all these altmetric platforms will be compared with that of highly cited articles to yield new insights into the differences and similarities between citations and altmetrics.

RESEARCH QUESTIONS

This study aims to examine the characteristics of highly mentioned titles on six altmetric platforms—Twitter, Weibo, Facebook, Wikipedia, blogs, and news—and compare them with that of highly cited titles. The motivations for this study are to detect differences and similarities between the six altmetric platforms and with citation events and to help to develop a clearer understanding of the different altmetric platforms. To reach these goals, this study examines two questions:

1. Do the characteristics of highly mentioned titles differ from those of highly cited titles?
2. Do the characteristics of the titles examined differ between each altmetric platform?

METHODS

For this work, highly mentioned titles on six altmetric platforms—Twitter, Weibo, Facebook, Wikipedia, blogs, and news— as identified by examining Altmetric.com data and highly cited articles from the Essential Science Indicators (ESI) database of Thomson Reuters were considered. A sample of 200 titles from each platform (n=1400) were selected representing the highest activity in each platform (i.e. highest number of tweets in Twitter, highest mentions on Weibo, highest mentions on Facebook, highest mentions in blogs and news outlets, highest number of appearances in Wikipedia articles, and highest number of citations in a specific field in ESI). Weibo data from Altmetric.com was collected from April 2014 through July 2015; Altmetric.com deprecated the Sina Weibo data source as of July 2015. Altmetric.com searches for assorted types of unique identifiers associated with academic research in various online sources and collects information about when, where and by whom the research is mentioned.

A codebook of article title types was developed based on a thorough review of the literature in combination with a grounded theory approach. The codebook included 11 variables that were identified, of which three (number of substantive words, total words, and attention grabbing

words) were discovered programmatically. Substantive words and total title words were programmatically identified using the Web VP Classic v.4 online tool (Nagano, 2015)¹, whereas the attention grabbing words (i.e., “New”, “Improved”, “Advanced”, “Superior”, “Beyond”, and “Better” (Subotic & Mukherjee, 2014)) were found using the *find* functionality in Excel. The coding of the 1,600 titles by an initial coder was completed in January 2016.

After discussion and coder training, a second coder chose the top 50 titles from each platform (n=200) and coded the titles based on six of the most subjective categories (declarative title, descriptive title, humor, references, named places, and acronyms). The second coder completed the coding in February 2016. The declarative title variable asked the coder to distinguish whether or not the title contained the conclusion(s) of the paper. Descriptive titles were coded if the title was neutral (did not contain conclusions). A paper title could not be both a declarative and descriptive title in the coding scheme. Humorous titles were coded when the title was humorous or light. The references coding included any mention of known entities outside of academia including proverbs, idioms, fictional characters, books, or music. With regards to named places, any mention of a specific location or where the study took place was captured. Finally, if an acronym was found in the title then the acronym was captured and the title was coded.

A Cohen’s kappa analysis was performed to determine inter-rater reliability for two of the subjective categories: (1) declarative and (2) humor. Inter-rater reliability was not tested for the three categories references, named places, and acronyms because the findings were small. In addition, interrater reliability was not tested on the descriptive category as the results would be identical to the declarative results (as it was an either/or choice). Results for the Cohen’s kappa for declarative (and descriptive) was quite substantial, as Kappa = 0.832 ($p < 0.001$). With regards to humor, Cohen’s kappa was found to be moderate to substantial, with Kappa = 0.73 ($p < 0.001$). These results indicate that there was high agreement for the declarative/descriptive categorization, and moderately high agreement for the humor categorization.

RESULTS

The title types of articles were examined across Twitter, Weibo, Facebook, Wikipedia, blogs, and news outlets and then compared with those titles from highly cited articles.

Table 1. Title types across different platforms

Title codes	Altmetric platforms						Citation
	Twitter	Weibo	Facebook	Blog	News	Wikipedia	
Number of special characters	81	85	69	58	41	62	77
Descriptive/ neutral title (%)	78	81.5	63	77.5	74	95.5	97
Declarative title (%)	21.5	17	37	22	26	4.5	3
Substantative words (%)	36.91	41.76	36.61	41.69	43.35	49.78	52.1
Title length	9.5	10.6	11.67	10.35	11.06	11.93	9.65
Geographically named places (%)	15.5	11	11	9	20	6.5	4.5
References outside/ known reference (proverbs, characters, books) (%)	8	1	4	3.5	1	0	0
Humor/ lightness (%)	6.5	4	4.5	2.5	2	0.5	0

¹ <http://www.lex Tutor.ca/vp/eng/>

Acronyms	6	10.5	6	8	4.5	39	24
Attention grabbing titles (%)	2.5	3	2	3	2	11	0

The results showed that highly cited articles on Wikipedia had an average length of 11.93 and were longer than titles of articles both on other altmetric platforms and from highly cited articles. Document titles from Facebook, news, Weibo, and blogs were longer than those found in Twitter and highly cited articles, respectively. The percentage of identified substantive words in the document titles indicate that highly cited titles had the highest percentage of substantive words (an average of 52.1%), whereas titles from all examined altmetric platforms contained an average of less than 50% substantive words. Regarding altmetric platforms, titles found on Wikipedia had the highest percentage of substantive words, whereas the titles from Facebook and Twitter had the lowest percentage, respectively. Document titles from Twitter and Weibo had the highest number of special characters and were followed by highly cited articles, whereas news and blog titles had the lowest number of characters. Except for document titles found on Wikipedia, titles from all other altmetric platforms were found to use mostly colon and question marks. The most prominent special characters found in highly cited titles were the colon, parenthesis, and hyphen.

The document titles were also coded as either containing research conclusions (declarative) or just a description of the research (descriptive). Although there were few titles that were neither declarative nor descriptive, most of the titles belonged to one of these two categories. The results showed that more than 60% of titles were coded as descriptive rather than declarative on all platforms. Approximately 3% to 4% of highly cited and Wikipedia titles were coded as declarative, while approximately 96% to 97% of the titles on the three platforms were descriptive. The findings indicate that Facebook had the highest percentage of declarative titles and the lowest percentage of descriptive titles among the platforms (37% and 63%, respectively). In addition, 26.5% of news and 22% blog titles were found to be declarative, while 74% and 77.5% respectively were descriptive. Of the highly tweeted titles, it was found that 21.5% were declarative and that 78% were descriptive. Weibo was the only platform where titles (2%) were found to be neither declarative nor descriptive; it was found that 17% of titles were declarative, while 81% were descriptive.

There were no titles providing humor or lightness coded among highly cited titles. The highest percentage of humorous/light titles were found on Twitter, where 6.5% of titles were coded as humorous. Approximately 4% of highly posted titles on Facebook and Weibo and 2% of news and blog titles were found to be light/humorous. Only 0.5% of highly posted titles on Wikipedia were coded as being humorous.

Highly cited and Wikipedia titles did not present any outside references, while 8% of Twitter and approximately 4% of Facebook and blog titles referenced outside material. Only 1% of news and Weibo titles referenced outside.

20% of news titles, 15.5% of tweeted titles, 11% of Weibo and Facebook titles, 9% of blog titles, 6.5% of Wikipedia titles, and 4.5% of highly cited titles presented geographical names (Table 1).

The results of One-Way ANOVA indicate that the title type difference between platforms is significant (Table 2).

Table 2. The results of One-way ANOVA for the difference between platforms in terms of different title codes in Table 1

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	61520.473	6	6835.61	91.776	0.000
Within Groups	5208.091	70	74.42		
Total	66728.564	79			

DISCUSSION

This study highlights the differences in document title types across seven contexts: Twitter, Weibo, Facebook, blog, news, Wikipedia, and citations. The percentage of substantive words in the titles indicated that highly cited titles had the highest percentage of substantive words (an average of 52.1%), whereas titles from the altmetric platforms contained an average of less than 50% substantive words. Among altmetric platforms, Wikipedia contained titles with the most substantive words, as compared to Twitter and Facebook. Substantive words in this scenario were counted in order to separate “functional from content words” (Nagano, 2015). The results suggest that because the use of substantive words (i.e., nouns)—perhaps considered in this case as scientific terms or jargon—occurred more often in titles found on Wikipedia and in highly cited documents, then the more complex and difficult to understand titles were more frequently read or referenced in these platforms and more likely to be cited. Wikipedia, Facebook, and news titles had the longest titles, while highly tweeted and highly cited titles were the shortest.

Descriptive titles were more visible than the declarative titles on all platforms. Jamali and Nikzad (2011) also showed that descriptive titles were more downloaded and cited than the declarative titles. The highest percentage of descriptive titles were found among highly cited documents in which 97% of titles were descriptive and only 3% were declarative. In addition, approximately 96% of highly mentioned titles on Wikipedia were descriptive, while the remaining were declarative. Weibo had 17% declarative titles and 81% descriptive titles. Titles of Facebook, news, blogs, and Twitter had between 63% to 78% descriptive titles, respectively.

No humorous/light titles were found among highly cited documents, while 6.5% of highly tweeted titles were coded as being humorous. Only 0.5% of Wikipedia titles were found to be humorous, whereas Facebook and Weibo had 4.5% and 4% humorous titles, respectively. The findings clearly show title type differences between general social media platforms (i.e. Twitter, Facebook and Weibo) and scholarly platforms (i.e. citations). The social media platforms are known for having a large population of public audiences while citation platforms are mainly used by scholars. This audience difference has certainly affected the visibility of documents and their title characteristics on these platforms. Wikipedia titles were found to be similar to highly cited titles, which could be because Wikipedia also has a large number of scholars as users and editors. A relevant study found that titles with a high level of amusement have a higher number of downloads, but that this does not correlate with the number of citations to the articles (Subotic & Mukherjee, 2014).

None of the highly cited titles and highly mentioned titles on Wikipedia referenced outside entities. Twitter had the highest percentage of titles referencing outside entities, including mentions of Kardashian and James Bond.

Only 4.5%, and 6.5% of highly cited titles and highly mentioned titles on Wikipedia included geographical names, while highly mentioned titles in news and highly tweeted titles had the highest percentage of titles presenting a geographical name. None of the highly cited titles included attention grabbing words, while 11% of highly mentioned titles on Wikipedia had such words. Both highly cited titles and highly mentioned titles on Wikipedia had the highest percentage of acronyms. A small percentage of highly mentioned document titles on news, Facebook, and Twitter had acronyms.

CONCLUDING REMARKS

The results from this study indicate the following similarities between citations and Wikipedia: These contexts had the highest percentage of substantive words, the highest percentage of descriptive titles, the lowest percentage of declarative titles, very few titles referencing outside entities or that contained humor/lightness, and finally they had the highest percentage of titles including acronyms. However, document titles were found to be different on Wikipedia from citations based on title length and presence of acronyms. A few similarities were detected between Twitter and Facebook including the same percentage of substantive words and acronyms in their titles and that the highest percentage of titles coded as being light/humorous among all altmetric platforms. Weibo document titles were similar to Twitter and Facebook in that they had the highest percentage of humorous titles after Facebook and Twitter. The similarities found from this work between titles in highly cited documents, Wikipedia, Twitter, and Facebook introduce new findings that can help scholars have a better understanding of altmetrics.

REFERENCES

- Ayres, I. & Vars, F. E. (1999). Determinants of citations to articles in elite law reviews. In: *Conference on Interpreting Legal Citations*, Evanston, Illinois.
- Björneborn, L. & Ingwersen, P. (2004). Toward a basic framework for webometrics. *Journal of the American Society for Information Science and Technology*, 55(14), 1216–1227.
- Cronin, B. (1984). *The citation process. The role and significance of citations in scientific communication*. London: Taylor Graham.
- Cronin, B. (1996). Rates of return to citation. *Journal of Documentation*, 52(2), 188-197.
- Jamali, H. R., & Nikzad, M. (2011). Article title type and its relation with the number of downloads and citations. *Scientometrics*, 88(2), 653-661.
- Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. Chicago: University of Chicago Press.
- Nagano, R. L. (2015). Research article titles and disciplinary conventions: A corpus study of eight disciplines. *Journal of Academic Writing*, 5(1). 133-144.
- Stremersch, S. Verniers, I. & Verhoef, P. C. (2007). The quest for citations: drivers of article impact. *Journal of Marketing*, 71, 171-193.
- Subotic, S. & Mukherjee, B. (2014). Short and amusing: The relationship between title characteristics, downloads, and citations in psychology articles. *Journal of Information Science February*, 40(1), 115-124.



What makes papers visible on social media? An analysis of various document characteristics¹

Zohreh Zahedi*, Rodrigo Costas*, Vincent Larivière** and Stefanie Haustein**

* z.zahedi.2@cwts.leidenuniv.nl; rcostas@cwts.leidenuniv.nl

Centre for Science and Technology Studies (CWTS), Leiden University, Wassenaarseweg 62A, Leiden, 2333 AL (The Netherlands)

** vincent.lariviere@umontreal.ca; stefanie.haustein@umontreal.ca

École de bibliothéconomie et des sciences de l'information, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, H3C 3J7 (Canada)

ABSTRACT

In this study we have investigated the relationship between different document characteristics and the number of Mendeley readership counts, tweets, Facebook posts, mentions in blogs and mainstream media for 1.3 million papers published in journals covered by the Web of Science (WoS). It aims to demonstrate that how factors affecting various social media-based indicators differ from those influencing citations and which document types are more popular across different platforms. Our results highlight the heterogeneous nature of altmetrics, which encompasses different types of uses and user groups engaging with research on social media.

INTRODUCTION

Five years after the introduction of the term (Priem, et al. 2010), altmetric indicators can be found on most of major publishers platforms, and are increasingly used in research evaluation (Wilsdon et al., 2015). Although some factors such as document age (Thelwall, et al, 2013), discipline (Haustein, Costas, & Larivière, 2015), topic (Costas, Zahedi, & Wouters, 2015) as well as countries (Alperin, 2015), have been shown to affect the various indicators, the processes which make a scientific paper visible on social and mainstream media are still not yet fully understood. Haustein et al. (2015) showed that factors which typically influence citations counts had a smaller or opposite effect on social and mainstream media mentions and that the usage pattern differed in particularly regarding document types. This study builds upon this work, taking into account a longer citation and social media window and expanding it by Mendeley readership counts. It addresses the following research questions:

What is the effect of document characteristics on the number of Twitter, Facebook, blogs and mainstream media mentions as well as on Mendeley readership counts? Particularly,

1. How do these effects compare with that observed for citations?
2. How do these effects differ across document types?

¹ This work was supported by the Alfred P. Sloan Foundation Grant #2014-3-25, Leiden University Fund (LUF) Grant # 4509/22-1-14/T, Vw and by funding from the DST-NRF Centre of Excellence in Scientometrics and STI Policy (South Africa).

DATA AND METHODS

This study builds upon and expands the analysis by Haustein et al. (2015) and compares the number of tweets, public Facebook posts, mentions in blogs and mainstream media, and Mendeley readership counts to citations received by WoS publications with a DOI published in 2012 ($n=1,339,279$). Citations from the CWTS in-house database were considered until September 2015 and altmetrics were collected in July 2015, expanding the windows used by Haustein et al. (2015). Twitter [T], Facebook [F], blogs [B] and mainstream media [M] mentions were obtained from Altmetric.com and Mendeley readership counts [MR] were collected using the Mendeley REST API.

The analyzed document properties included the document type as indicated by WoS [DT], the number of pages [PG], cited sources in the reference list (including non-source items) [NR], and characters in the title [TI], as well as number of authors [AU], institutions [IN] and countries [CU] of the paper. The percentage of papers with at least one citation or social and mainstream media event count (coverage), the average number of counts per paper (density) and the average number of counts for documents with at least one count (intensity) were computed. Correlations are based on Spearman's ρ .

RESULTS AND DISCUSSION

Table 1 shows that slightly more papers had been saved to Mendeley (84.2%) than cited (81.7%). For other social media platforms, coverage is much lower, with 22.6% of papers receiving at least one tweet, 5.2% being shared publicly on Facebook, 2.3% mentioned in blog posts, and 1.1% discussed by mainstream media. Reviews and articles are the document types that were most commonly cited or saved on Mendeley, while editorial material and news items were particularly popular on Twitter, Facebook, blogs, and mainstream media. Although both coverage and density were higher for reviews and articles, editorials and news items were also frequently saved by Mendeley users.

Table 1. Prevalence (coverage in %, density, intensity) of citations and social media metrics per document type.

		All document types	Article	Biographical Item	Book Review	Correction	Editorial Material	Letter	Meeting Abstract	News Item	Review
<i>N</i>		1,338,885	1,132,428	2,302	21,710	9,817	60,533	29,410	13,071	4,880	64,734
%		99.97	84.56	0.17	1.62	0.73	4.52	2.2	0.98	0.36	4.83
Citations	Coverage	81.72%	86.89%	9.86%	2.13%	49.53%	49.53%	47.49%	6.34%	37.34%	94.14%
	Density	7.68	7.84	0.17	0.03	0.44	2.26	1.69	0.09	1.39	18.78
	Intensity	9.40	9.02	1.70	1.23	2.31	4.57	3.56	1.35	3.72	19.95
Blogs	Coverage	2.28%	2.27%	0.70%	0.18%	2.21%	2.71%	1.00%	0.05%	2.68%	3.85%
	Density	0.04	0.04	0.01	0.00	0.03	0.05	0.02	0.00	0.03	0.06
	Intensity	1.78	1.80	1.06	1.10	1.21	1.77	1.78	1.00	1.27	1.62
Twitter	Coverage	22.55%	21.98%	13.99%	5.42%	10.68%	28.57%	19.02%	2.21%	47.97%	38.67%
	Density	1.02	0.94	0.41	0.11	0.20	2.05	0.59	0.05	4.26	1.95
	Intensity	4.52	4.30	2.90	1.98	1.90	7.17	3.12	2.46	8.89	5.05
Facebook	Coverage	5.20%	4.94%	3.00%	1.31%	1.33%	7.99%	3.64%	0.27%	11.13%	7.99%

	Density	0.12	0.11	0.00	0.02	0.02	0.18	0.07	0.00	0.21	0.25
	Intensity	2.22	2.21	1.45	1.21	1.20	2.28	1.98	1.11	1.90	2.46
Mainstream media	Coverage	1.08%	1.10%	0.13%	0.00%	0.12%	1.15%	0.31%	0.01%	0.68%	1.67%
	Density	0.02	0.02	0.00	0.00	0.00	0.02	0.01	0.00	0.01	0.03
	Intensity	2.02	2.05	1.00	2.00	1.50	1.97	3.14	1.11	1.91	1.66
Mendeley readership	Coverage	84.23%	87.70%	28.37%	28.83%	25.37%	68.23%	58.36%	24.74%	59.41%	93.57%
	Density	11.00	10.94	0.79	1.51	3.41	6.47	2.76	0.57	4.98	27.29
	Intensity	13.06	12.47	2.80	5.23	13.46	9.48	4.73	2.31	8.38	29.16

Correlations show that Mendeley readership has the highest positive correlation ($\rho = .585$, Table 2) with citation counts, followed by Twitter ($\rho = .279$) and blogs ($\rho = .159$), while Facebook ($\rho = .142$) and mainstream media ($\rho = .115$) show positive but low correlations with citations. These findings point to different audiences and engagements on these social media platforms. While the stronger relationship between citations and readership counts likely reflect Mendeley's use in a pre-citation context (Mohammadi, Thelwall, & Kousha, 2015), the lower correlations with Twitter might be related to Twitter's inclusion of non-academic audiences. Facebook is mostly used for private rather than professional purposes (Van Noorden, 2014), and users generally interact in closed rather than open groups. Blogs and mainstream media are very selective in the sense that only a fraction of papers are mentioned. It should be noted that the low correlations are largely caused by low coverage: more than 98% of papers did not get mentioned in blogs or mainstream media. Both of these sources are targeted at larger audiences than scientific papers and are generally written in a less technical language, while blogs mainly focus on academia and mainstream media target a general audience. It should also be mentioned that papers covered by mainstream media and blogs are often published in multidisciplinary scientific journals such as *Nature* or *Science* (Costas et al, 2015).

Table 2. Correlation between document characteristics, citations and social media mentions.

	PG	NR	TI	AU	IN	CU	C	MR	B	T	F	M
PG	1.000	0.622	0.079	-0.006	0.116	0.131	0.250	0.287	0.007	0.036	0.013	-0.001
NR		1.000	0.165	0.155	0.168	0.146	0.485	0.471	0.061	0.145	0.068	0.043
TI			1.000	0.323	0.135	0.038	0.169	0.080	-0.033	-0.007	-0.012	-0.022
AU				1.000	0.494	0.252	0.320	0.168	0.031	0.085	0.047	0.033
IN					1.000	0.560	0.215	0.177	0.049	0.102	0.061	0.042
CU						1.000	0.170	0.153	0.045	0.060	0.039	0.036
C							1.000	0.585	0.140	0.220	0.120	0.108
MR								1.000	0.159	0.279	0.142	0.115
B									1.000	0.211	0.193	0.297
TW										1.000	0.328	0.161
FB											1.000	0.182
M												1.000

(N=1,339,279); PG=Page, NR= Number of References, TI= Title length, AU=Author, IN=Institute, CU=Country; C=Citations, MR=Mendeley readership counts, B=Blogs, T=Twitter, F=Facebook, M=Mainstream Media

At the level of document characteristics, Mendeley readership counts exhibit the highest positive correlation with the number of references made ($\rho = .471$) — showing trends similar

to citations — followed by number of pages ($\rho=.287$) and title length ($\rho=.080$). However, the latter does not seem to have a large effect on attracting Mendeley users. Readership count patterns are comparable to citations. Although correlations were low, negative correlations between the other metrics and title length (as well as document length for main stream media) suggests that social media users, to the opposite of citing authors, exhibit a preference for short titles (and documents length). The highest correlation of citations is with the number of authors ($\rho=0.320$), followed by number of institutions ($\rho=0.215$) and countries ($\rho=0.170$). Altmetrics show less pronounced effects regarding these collaboration indicators slightly different and less pronounced effects.

PRELIMINARY CONCLUSIONS AND OUTLOOK

This paper provided insights on the relationship between social and mainstream media visibility and various documents characteristics. It is shown that some of them influence the extent to which they are cited or shared on social media. However, patterns vary between indicators. While Twitter, Facebook, blogs and mainstream media mentions are different from citations as reflected in low correlations and the popularity of so-called “non-citable” document types, Mendeley exhibits patterns similar to citations, which is likely due to its use in a pre-citation context. Our results thus highlight the heterogeneous nature of altmetrics, which encompasses different types of uses and user groups engaging with research on social media. Future research will include to what extent this pattern is different across disciplines as well as how these indicators change by different levels of collaboration and document characteristics by applying multiple regression analysis.

REFERENCES

- Alperin, J. P. (2015). Geographic variation in social media metrics: an analysis of Latin American journal articles. *Aslib Journal of Information Management*, 67(3), 289–304. <http://doi.org/10.1108/AJIM-12-2014-0176>
- Costas, R., Zahedi, Z., & Wouters, P. (2015). The thematic orientation of publications mentioned on social media: Large-scale disciplinary comparison of social media metrics with citations. *Aslib Journal of Information Management*, 67(3), 260–288. <http://doi.org/10.1108/AJIM-12-2014-0173>
- Haustein, S., Costas, R., & Larivière, V. (2015). Characterizing social media metrics of scholarly papers: The effect of document properties and collaboration patterns. *PLoS ONE*, 10(3), e0120495. <http://doi.org/10.1371/journal.pone.0120495>
- Mohammadi, E., Thelwall, M., & Kousha, K. (2015). Can Mendeley Bookmarks Reflect Readership? A Survey of User Motivations. *Journal of the Association for Information Science and Technology*, 67(5). <http://doi.org/10.1002/asi.23477>
- Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010, October 26). Altmetrics: A manifesto. Retrieved from <http://altmetrics.org/manifesto/>
- Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do Altmetrics Work? Twitter and Ten Other Social Web Services. *PLoS ONE*, 8(5), e64841. <http://doi.org/10.1371/journal.pone.0064841>
- Van Noorden, R. (2014). Online collaboration: Scientists and the social network. *Nature*, 512(7513), 126–129. <http://doi.org/10.1038/512126a>



Normalization of Mendeley reader impact on the reader- and paper-side ¹

Robin Haunschild* and Lutz Bornmann**

* R.Haunschild@fkf.mpg.de

Max Planck Institute for Solid State Research, Heisenbergstr. 1, Stuttgart, 70569 (Germany)

** bornmann@gv.mpg.de

Division for Science and Innovation Studies, Administrative Headquarters of the Max Planck Society, Hofgartenstr. 8, Munich, 80539 (Germany).

ABSTRACT

For the normalization of citation counts, two different kinds of methods are possible and used in bibliometrics: the cited-side and citing-side normalizations both of which can also be applied in the normalization of “Mendeley reader counts”. Recently, we introduced the paper-side normalization of reader counts (mean normalized reader score, MNRS) which is an adaptation of the cited-side normalization. Since the calculation of the MNRS needs further data besides data from Mendeley (a field-classification scheme, such as the Web of Science subject categories), we introduce here the reader-side normalization of reader counts which is an adaptation of the citing-side normalization and does not need further data from other sources. In this study, all articles and reviews of the Web of Science core collection with publication year 2012 (and a DOI) are used to normalize their Mendeley reader counts. The newly proposed indicator (mean discipline normalized reader score, MDNRS) is obtained, compared with the MNRS and bare reader counts, and studied empirically. We find that: (i) normalization of Mendeley reader counts is necessary, (ii) the MDNRS is able to normalize Mendeley reader counts in several disciplines, and (iii) the MNRS is able to normalize Mendeley reader counts in all disciplines. This generally favorable result for the MNRS in all disciplines lead to the recommendation to prefer the MNRS over the MDNRS – provided that the user has an external field-classification scheme at hand.

INTRODUCTION

Normalization of citation counts regarding subject category and publication year of publications started in the mid-1980s. The comparison of units in research (e.g. researchers, research groups, institutions, or countries) publishing in different disciplines and time periods is only possible with normalized citation scores. Basically, one can distinguish between two levels of normalization:

(1) In the case of normalization on the cited side, the total number of citations of the paper *i* to be evaluated is counted. This number of times cited is compared with other publications published in the same year and subject category (the reference set). The mean citation rate over the papers in the reference set determines the expected value. The comparison of times cited of paper *i* with the expected value results in the normalized citation score (NCS) for *i*. This procedure is repeated for all papers in a paper set (e.g. of a researcher, research group,

¹ This work was supported by Mendeley and the Max Planck Digital Library. It is based on <https://dx.doi.org/10.6084/m9.figshare.2554957.v1>. A full and extended paper has been published in the *Journal of Informetrics* (Bornmann & Haunschild, 2016).

institution, or country) and a mean NCS (MNCS) on a higher aggregation level is calculated (Waltman, et al., 2011).

(2) In the case of normalization on the citing side, each citation of a paper is multiplied with a weighting factor (Zitt & Small, 2008). This weighting factor reflects the citation density of the discipline: Since it is assumed that the number of references in publications reflects the field-specific citation density, the inverse of the field-specific citation density is usually used as a weighting factor. The sum of all weighted citations is the normalized citation impact of a publication. The results of Waltman and van Eck (2013) offer considerable support for the use of citing-side indicators.

In recent years, scientometrics started to explore alternative metrics (altmetrics) to study the impact of publications (Priem, 2014). Here, notes, saves, tweets, shares, likes, recommends, tags, posts, trackbacks, discussions, bookmarks, comments etc. are counted (Bornmann, 2014a). Altmetrics seems to have two advantages over citation counts: They allow (1) an impact measurement within a shorter time period after the appearance of a paper than citation counts and (2) a broader impact measurement, which is not only restricted to the area of science but also to the rest of society. The possibility of a broad impact measurement using altmetric counts is part of the current scientometric research (Bornmann, 2014b).

Data from Mendeley (Elsevier) which reflect the readership of papers are one of the most important sources for altmetrics. Mendeley is both a citation management tool and social network for academics. One basic assumption for using Mendeley data as altmetrics is that Mendeley users who add publications to their libraries can be counted as readers of the publications. Haunschild and Bornmann (2016) introduced the paper-side normalization of “reader counts”, because several studies have shown that Mendeley reader impact – similar as citation impact – varies across scientific disciplines (see e.g. Haustein & Larivière, 2014a; Zahedi & Eck, 2014): In one discipline papers are read more often on average than in other disciplines. The new indicator on the paper-side was named mean normalized reader score (MNRS). The normalization also considers the document type of publications, because it has an influence on reader counts, too (Haustein & Larivière, 2014b). Since citing-side normalization is a promising alternative in scientometrics to cited-side normalization and the MNRS needs further data besides data from Mendeley, we introduce the mean discipline normalized reader score (MDNRS) in this study. The MDNRS is an adaptation of the citing-side normalization to reader data and is solely based on Mendeley data.

METHODS

It is common practice in bibliometrics to include only articles and reviews in a study. We retrieved the Mendeley reader statistics for articles and reviews published in 2012 and having a DOI ($n_A = 1,133,224$ articles and $n_R = 64,960$ reviews). The DOIs of the papers from 2012 were exported from the in-house database of the Max Planck Society (MPG) based on the WoS and administered by the Max Planck Digital Library (MPDL). We used R (<http://www.r-project.org>) to interface to the Mendeley API. We used the Mendeley API which was made available in 2014. DOIs were used to identify the papers in the Mendeley API. We found 1,074,407 articles (94.8%) and 62,771 reviews (96.6%) at Mendeley.

In total, the articles were registered 9,347,500 times and the reviews were registered 1,335,233 times with a sub-discipline. The sub-disciplines are self-assigned by the Mendeley users. Only 4,924 (0.05%) of the Mendeley article readers and 531 (0.04%) review readers did not declare any discipline information. For 118,167 articles (10.4%) and 4,348 reviews (6.7%) we found the paper at Mendeley but without a reader. In total, 956,105 articles with 9,347,500 reader counts (approximately 10 readers per article) and 58,420 reviews with 1,335,233 reader counts (approximately 23 readers per article) were used in this study. The papers without any reader are not used in the normalization procedure introduced here (the MDNRS), because it is not

possible to identify a discipline for a reader count which was not declared by the user. However, papers without any reader and papers not found at Mendeley were used for the calculation of the MNRS as zero-reader papers which is included in this study for the comparison with the MDNRS.

The requests to the Mendeley API were made from December 11-23, 2014. All data in this study are based on a partial copy of our in-house database (last updated on November 23, 2014) supplemented with the Mendeley reader counts.

RESULTS

Differences in reader impact between disciplines

The Mendeley reader counts broken down by discipline are shown in Table 1. 95.5% of the readers are within 15 of the 25 disciplines while the remaining 4.5% of the readers are in the 10 disciplines with less than 1% each. The disciplines biological sciences and medicine comprise 48% of the Mendeley readers of the WoS papers from 2012.

Table 1. Basic variables to construct the MDNRS: Mendeley reader counts of WoS papers (articles and reviews) from 2012 and average number of Mendeley readers broken down by the different Mendeley disciplines (sorted in decreasing order of the raw Mendeley reader counts)

Mendeley discipline	Mendeley reader counts		Average number of readers	
	abs. readers	% readers	Reviews	Articles
Biological Sciences	3,518,931	32.94	14.30	6.85
Medicine	1,610,631	15.08	6.81	3.84
Chemistry	852,261	7.98	5.99	3.42
Engineering	709,525	6.64	4.66	3.26
Physics	578,831	5.42	4.35	3.46
Psychology	567,297	5.31	7.41	4.92
Environmental Sciences	406,960	3.81	4.42	3.81
Computer and Information Science	363,337	3.40	2.68	2.77
Social Sciences	354,877	3.32	2.54	2.76
Earth Sciences	319,943	2.99	1.78	2.58
Materials Science	289,464	2.71	4.13	2.39
Electrical and Electronic Engineering	194,604	1.82	4.24	2.62
Business Administration	173,815	1.63	3.11	3.52
Economics	133,370	1.25	1.86	2.17
Education	133,026	1.25	2.67	2.26
Management Science	91,340	0.86	1.95	2.02
Astronomy, Astrophysics, and Space Science	80,713	0.76	3.06	2.89
Mathematics	77,496	0.73	1.55	1.54
Sports and Recreation	54,699	0.51	2.36	2.16
Humanities	45,094	0.42	1.31	1.38
Design	35,935	0.34	1.33	1.22
Arts and Literature	30,756	0.29	1.14	1.15
Linguistics	27,162	0.25	2.33	2.32
Philosophy	21,121	0.20	1.32	1.47
Law	11,545	0.11	1.29	1.3

The average number of Mendeley readers per paper is also shown in Table 1 because this is the basic variable to construct the new indicator. The results reveal that the average number of readers varies significantly across the Mendeley disciplines. Furthermore, the average number of Mendeley readers per paper is larger for reviews than for articles in 15 of the 25 disciplines.

Reader-side normalization of reader impact

As Table 1 shows that the readers differ on average between the document types “article” and “review” for most Mendeley disciplines, the normalization procedure is done separately for both document types.

The procedure for normalizing reader counts on the reader side is as follows: First, the average number of reader counts in each Mendeley discipline (ρ_d) is determined (see Table 1):

$$\rho_d = \frac{1}{I_d} \sum_{i=1}^{I_d} R_{id}$$

(4)

Here, I_d is the number of papers in Mendeley discipline d and R_{id} is the raw Mendeley reader count of paper i in Mendeley discipline d . A paper is in discipline d if at least one of its readers is in this discipline. The average reader count (ρ_d) should reflect differences between disciplines in reading papers.

Second, the Mendeley reader counts (R_{id}) of paper i and Mendeley discipline d is divided by the average reader count (ρ_d) in discipline d :

$$\beta_{id} = \frac{R_{id}}{\rho_d}$$

(5)

Third, the sum over the normalized reader counts (β_{id}) in the disciplines in which a paper i was read is calculated:

$$\text{DNRS}_i = \sum_{d=1}^D \beta_{id}$$

(6)

In Eq. (6), D is the number of Mendeley disciplines (currently $D = 25$ at Mendeley). We obtain a normalized reader score (discipline normalized reader score for each paper i , DNRS_i). Similar to the citing-side normalization of citation counts, where each citation is weighted by the citation density in a discipline (reflected by the number of references), each reader of a publication is weighted by the corresponding reader density (reflected by the average readers in a discipline). Since reader counts are dependent on time (the longer the time window between the publication of a paper and its impact measurement, the more readers can be expected), the DNRS_i should be calculated separately for papers published in different years (Lin & Fenner, 2013).

The overall reader impact for aggregation levels (e.g. single researchers, research groups, institutions, countries, or journals) can be analyzed in terms of averages over paper sets:

$$\text{MDNRS} = \frac{1}{N} \sum_{i=1}^N \text{DNRS}_i$$

(7)

Empirical analysis of the discipline normalized reader score

Bornmann, de Moya Anegón, and Mutz (2013) introduced a statistical procedure which can be used to study the proposed ability of the MDNRS to field-normalize. In order to compare the results for the MDNRS with other reader indicators, the procedure is also applied to the MNRS and bare reader counts. In the first step of the procedure (done for each indicator separately), all papers from 2012 are sorted in descending order by an indicator. Then, the 10% most frequently read papers are identified (a new binary variable is generated).

In the second step, the papers are assigned to main disciplines using the OECD field classification scheme. The main OECD disciplines aggregate WoS subject categories which consist of sets of disciplinary journals to the following broad disciplines: (1) natural sciences, (2) engineering and technology, (3) medical and health sciences, (4) agricultural sciences, (5) social sciences, and (6) humanities.

In the third step, the proportion of papers which belong to the 10% most frequently read papers from the first step is determined for each broad discipline. The expectation is that this proportion equals 10% if the indicator values are independent of disciplines or are properly field-normalized, respectively. Thus, bare reader counts should show greater deviations from 10% than MNRS and MDNRS.

Table 2. Number of papers and proportion of papers belonging to the 10% most frequently read papers in six main disciplines (as defined by the OECD)

Main disciplines	Bare reader counts		MDNRS		MNRS	
	Number of papers	Proportion top-10%	Number of papers	Proportion top-10%	Number of papers	Proportion top-10%
Natural Sciences	703,380	12.3	703,380	11.3	830,635	10.1
Engineering and technology	318,496	7.4	318,496	11.0	382,496	10.6
Medical and health sciences	440,094	8.3	440,094	6.9	506,721	9.6
Agricultural sciences	52,527	6.9	52,527	5.4	59,108	9.8
Social Sciences	129,565	17.9	129,565	25.0	137,551	9.2
Humanities	17,506	4.6	17,506	10.4	24,378	10.3
Mean deviation		3.8		4.2		0.4

The results of the three-step procedure are shown in Table 2. The MDNRS is compared with the MNRS and bare reader counts. The table shows the total number of papers within the main disciplines and the proportion of papers within a main discipline which belongs to the 10% most frequently cited papers. As the number of papers for bare reader counts, MDNRS, and MNRS show, the paper numbers for MNRS are significantly higher in all main disciplines. This is due to the fact that papers with zero readers and papers from 2012, which could not be found on Mendeley, cannot be considered for the analyses of the bare reader counts and MDNRS.

The MNRS shows the best results in Table 2: All main disciplines have less than 1 percent point deviations from 10%. Following the argumentations of Sirtes (2012) and Waltman and van Eck (2013), the comparably best results for the MNRS could have a simple reason: The indicator uses the same scheme of subject categories for the field-normalization on which the tests in Table 2 are also based. Waltman and van Eck (2013) therefore repeated the analyses using another scheme of field categorization: an algorithmically constructed classification system

(ACCS). The ACCS is based on direct citation relations between publications. The results of the comparison between bare reader counts, MDNRS, and MNRS based on ACCS (applied on the highest field-aggregation level) is shown in Table 3.

Table 3. Number of papers and proportion of papers belonging to the 10% most frequently read papers in five main disciplines (as defined by the ACCS on the highest level)

Main disciplines	Bare reader counts		MDNRS		MNRS	
	Number of papers	Proportion top-10%	Number of papers	Proportion top-10%	Number of papers	Proportion top-10%
Biomedical and health sciences	476,324	10.6	476,324	7.9	859,504	9.5
Life and earth sciences	205,282	14.4	205,282	11.0	383,988	11.1
Mathematics and computer science	83,412	5.6	83,412	10.4	207,605	9.3
Physical sciences and engineering	326,582	6.6	326,582	8.9	711,010	10.0
Social sciences and humanities	113,710	16.2	113,710	22.2	206,633	11.9
Mean deviation		3.8		3.4		0.8

The results are similar to those in Table 2. The MNRS reveals the best result: The proportions of papers belonging to the 10% most frequently read papers fall within the range of tolerance in all disciplines. MDNRS follows with two out of five disciplines having larger deviations from the expected value. However, there is with 22.2% a large deviation for the social sciences and humanities (similar to Table 2).

Taken as a whole, the proportions in Table 2 and Table 3 for bare reader counts reveal that field-normalization is generally necessary for Mendeley reader counts. Larger deviations from the expected value of 10% are found in most of the disciplines. The MNRS should be preferred for the field-normalization, because it seems to reach the desired goal in all disciplines.

DISCUSSION

Here, we have proposed the field-normalized indicator MDNRS based on Mendeley data which might complement the paper-side normalization of reader counts (MNRS). Since the calculation of the MNRS needs further data besides data from Mendeley (in order to have a field-classification scheme for normalization), the reader-side normalization of reader counts is an attractive alternative: The MDNRS does not need further data and can be exclusively calculated with data from Mendeley, because the MDNRS is normalized with respect to Mendeley disciplines which are reported with virtually all reader counts.

In this study, we tested whether the MDNRS is able to field-normalize reader counts. For comparison, we included also bare reader counts and the MNRS in the analyses. The MNRS shows the best results in general and can be recommended as a properly field normalized indicator in all disciplines. The results for the MDNRS are ambivalent, whereby the social sciences are the most problematic discipline with large deviations from the expected value. However, both methods of field-normalization received significantly better results than bare reader counts.

REFERENCES

- Bornmann, L. (2014a). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics*, 8(4), 895-903. doi: 10.1016/j.joi.2014.09.005.
- Bornmann, L. (2014b). Validity of altmetrics data for measuring societal impact: A study using data from Altmeteric and F1000Prime. *Journal of Informetrics*, 8(4), 935-950.
- Bornmann, L., de Moya Anegón, F., & Mutz, R. (2013). Do universities or research institutions with a specific subject profile have an advantage or a disadvantage in institutional rankings? A latent class analysis with data from the SCImago ranking. *Journal of the American Society for Information Science and Technology*, 64(11), 2310-2316.
- Bornmann, L. & Haunschild, R. (2016) Normalization of Mendeley reader impact on the reader- and paper-side: A comparison of the mean discipline normalized reader score (MDNRS) with the mean normalized reader score (MNRS) and bare reader counts. *Journal of Informetrics*, 10(3), 776-788.
- Haunschild, R. & Bornmann, L. (2016). Normalization of Mendeley reader counts for impact assessment. *Journal of Informetrics*, 10(1), 62-73.
- Haustein, S., & Larivière, V. (2014a). Mendeley as a Source of Readership by Students and Postdocs? Evaluating Article Usage by Academic Status. Paper presented at the Proceedings of the IATUL Conferences. Paper 2.
- Haustein, S., & Larivière, V. (2014b). A multidimensional analysis of Aslib proceedings – using everything but the impact factor. *Aslib Journal of Information Management*, 66(4), 358-380. doi: doi:10.1108/AJIM-11-2013-0127.
- Lin, J., & Fenner, M. (2013). The many faces of article-level metrics. *Bulletin of the American Society for Information Science and Technology*, 39(4), 27-30. doi: 10.1002/bult.2013.1720390409.
- Priem, J. (2014). Altmetrics. In B. Cronin & C. R. Sugimoto (Eds.), *Beyond bibliometrics: harnessing multi-dimensional indicators of performance* (pp. 263-288). Cambridge, MA, USA: MIT Press.
- Sirtes, D. (2012). Finding the Easter eggs hidden by oneself: why fairness test for citation indicators is not fair. *Journal of Informetrics*, 6(3), 448-450. doi: 10.1016/j.joi.2012.01.008.
- Waltman, L., & van Eck, N. J. (2013). A systematic empirical comparison of different approaches for normalizing citation impact indicators. *Journal of Informetrics*, 7(4), 833–849.
- Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. J. (2011). Towards a new crown indicator: some theoretical considerations. *Journal of Informetrics*, 5(1), 37-47. doi: 10.1016/j.joi.2010.08.001.
- Zahedi, Z., & Eck, N. J. v. (2014). Visualizing readership activity of Mendeley users using VOSviewer. Retrieved July 8, 2014, from http://figshare.com/articles/Visualizing_readership_activity_of_Mendeley_users_using_VOSviewer/1041819
- Zitt, M., & Small, H. (2008). Modifying the journal impact factor by fractional citation weighting: The audience factor. *Journal of the American Society for Information Science and Technology*, 59(11), 1856-1860. doi: 10.1002/asi.20880.

CHAPTER 16

Citation Impact



Web of science coverage and scientific performance of Central and Eastern European countries¹

Adam Ploszaj*, Agnieszka Olechnicka**

* aploszaj@gmail.com

Centre for European Regional and Local Studies (EUROREG), University of Warsaw, Krakowskie Przedmieście str. 30, 00-927 Warsaw, Poland

** a.olechnicka@uw.edu.pl

Centre for European Regional and Local Studies (EUROREG), University of Warsaw, Krakowskie Przedmieście str. 30, 00-927 Warsaw, Poland

ABSTRACT

The paper put forward the hypothesis that changes in the list of journals, particularly in the number of periodicals published in specific Central and Eastern European (CEE) countries, have a significant impact on bibliometric indicators and, consequently, on the convergence levels they are used to measure. The analyses presented in the article support the argument that countries from Central and Eastern Europe, despite showing fairly consistent convergence trends, achieve noticeably weaker results than Western Europe regarding research and development and scientific activity. The evident impact made by the inclusion of numerous CEE journals in WoS on the values of the indicator analysed, directly supports the hypothesis put forward in this article. The results of the study are important because data on the number of publications and citations in the Web of Science are increasingly used as development indicators of national R&D sectors. By showing how modifications in these databases influence the results obtained, we can better understand and thus make better use of data from these sources. The article concludes with listing the possibilities for furthering and deepening selected themes pointed out in the paper.

INTRODUCTION

Our analysis serves to show a phenomenon that has up till now been largely overlooked: the way in which modifications to the coverage of bibliometric databases influence observable trends, in our case the process of scientific convergence of CEE countries with Western Europe. To conduct such an analysis, we make use of detailed data from the Web of Science database. The usual analyses conducted are based on the overall number of articles attributed to a given country. From this angle, the growth and convergence of CEE countries are as clear as day. However, this approach ignores the fact that the list of journals in the WoS database is not permanent but, quite the contrary, changes significantly in some periods. We put forward the hypothesis that changes in the list of journals, particularly in the number of periodicals published in specific CEE countries, have a significant impact on bibliometric indicators and, consequently, on the convergence levels they are used to measure. In a broader sense, our analysis aims to show that observed trends in scientific output sometimes result not only from intensified research activity but may also be the effect of elements being accounted for which

¹ This work was supported by the Polish National Science Center under grant “Polish scientific centres in the European cooperation network - characteristics, determinants, mechanisms” and European Union's 7th Framework Programme under project “GRINCOH—Growth-Innovation-Competitiveness: Fostering Cohesion in Central and Eastern Europe”.

were not previously included, in short, they derive from more precise (or simply different) measures. This viewpoint is also important because data on the number of publications and citations in the Web of Science are increasingly used as development indicators of national R&D sectors (EC 2014a). By showing how modifications in these databases influence the results obtained, we can better understand and thus make better use of data from these sources.

The spatial scope of the study was defined as the ten countries of Central and Eastern Europe (EU10) which acceded to the EU as part of expansion in 2004 and 2007. These are the Czech Republic, Estonia, Lithuania, Latvia, Poland, Slovakia, Slovenia, Hungary, Bulgaria and Romania. In the article their situation is outlined against the background of the so-called ‘old’ member states, also referred to as ‘the fifteen’ (EU15) as well as in reference to the situation of the whole European Union (EU27, i.e. without Croatia that entered the EU on 1 July 2013).

In the study, alongside widely available data concerning expenditure and employment in research and development, bibliometric data was used that was generated from the Web of Science database. Detailed data, on the level of individual articles, was retrieved from the database in June 2014. The sample created from this source contains in total 547,050 articles affiliated to Central and Eastern Europe (EU10) published in the years 2000-2013. These constitute 10.9% of articles affiliated to EU27 countries (5,034,893).

RESULTS

The Web of Science does not include all scientific articles published worldwide. It contains only articles from selected journals. In principle, the choice of journals is meant to be content-based – the intention is to include the most prominent (the best) periodicals. However, other factors also play a role. One of these is the desire to achieve a spatially balanced database, by including journals that are not only important on the world scale, but also those whose scope is more regional (i.e. a group of countries) or even national (Testa 2011). This approach is particularly important in the case of periodicals in the field of social science and humanities, as the research problems they deal with are often of a national, linguistic or even local nature. However, journals relating to the exact sciences which are clearly of a national character (supported by the fact that they are published in non-congress languages) can also be included in WoS. For example, it is a case of the journal “Przemysł Chemiczny” (Chemical Industry) which is published in Poland and mostly in the Polish language (there is a fraction of articles in English).

Inclusion in the database is not indefinite. If a journal does not fulfil the criteria determined by WoS, it is simply removed. The criteria include formal requirements (e.g. regular issue) as well as content requirements (a suitably high Impact Factor)². When a given journal is removed from the index, the ‘space’ becomes available to new titles. The scope for considering new titles depends both on the engagement of publishers who apply for entry, as well as of the database administrators who, apart from respecting the overriding principle of listing the most important journals, must to some extent be guided by business sense. We should also remember, that WoS is a commercial product, and its creators (owners) are guided by economic outcomes. It follows that, in aiming to include journals from different countries (markets), we can perceive that WoS does not only strive to provide an adequate representation of the diversity of worldwide research but also seeks to reach new potential

² <http://wokinfo.com/essays/journal-selection-process/>, access: 14.01.15.

clients. Moreover, an important question remains unanswered as to whether the Thomson Reuters' selection criteria are consistent and rigorous or whether some countries are favoured and overrepresented in WoS (Kosanovic and Sipka 2013).

In recent years we have witnessed a fairly important expansion in the spatial range of the WoS database. This growth has been markedly more intense than in previous decades. In the years 1980-1990, the number of journals in the database rose from 6130 by 654 titles (10.7%) In the decade 1990-2000, a further 1444 journals were added (21.3%). In the years 2000-2010 the increase was visibly steeper: in 2000 there were 8228 titles and as many as 11793 in 2010. The increase of 3511 journals meant that the list of titles grew by 42.7%. This radical expansion of the database is described by Thomson Reuters as "The Globalization of the Web of Science" (Testa 2011). Apart from the routine analysis of journals for inclusion in the database – in recent years around 2.5 thousand applications annually, of which around 10% are accepted – in the years 2007-2009 action was undertaken to increase the database's representation of journals outside the 'centre' of world research:

[...] from 2007 to 2009 the Editorial Development Department at Thomson Reuters focused on a collection of more than 10,000 regional journals (these are journals published outside the US or the UK that contain the scholarship of authors from a particular region or country, and cover topics of regional interest or topics studied from a regional perspective). Sixteen hundred (1,600) of these 10,000 journals met Thomson Reuters standards and were selected for coverage (Testa 2011, s. 2).

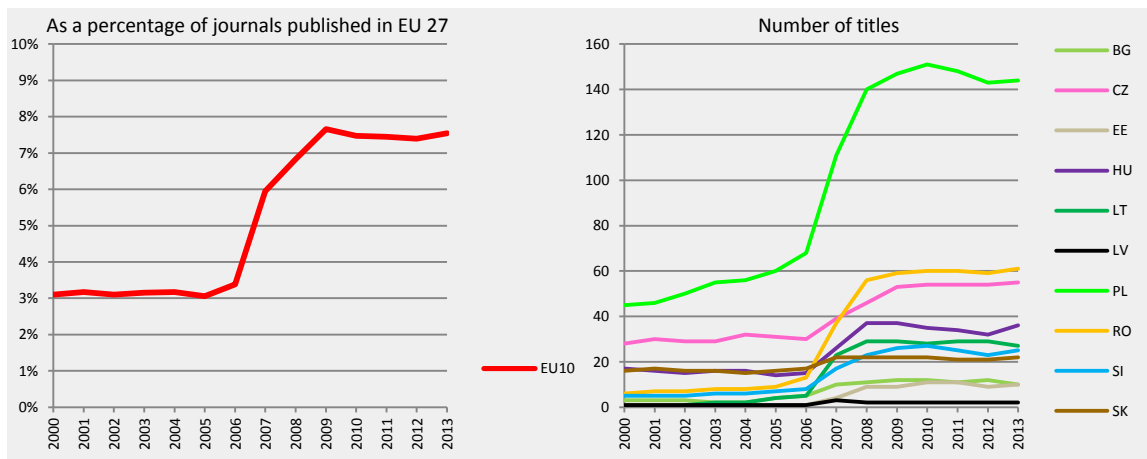
In consequence, the number of journals published in some countries and listed in WoS grew significantly between 2005 and 2010. The steepest growth in absolute numbers was recorded in: Spain (112 new titles; growth of 207%), Brazil (105; 389%), Australia (97; 105%), Poland (85; 149%), Turkey (68; 971%), Italy (68; 56%), France (62; 28%), China (62; 75%), South Korea (62; 168%), Japan (61; 35%), India (60; 113%), Romania (52; 650%), Croatia (47; 336%), South Africa (41; 152%) (Testa 2011, p. 4). The increase in the number of journals naturally leads to an increased number of indexed articles in WoS. This should be remembered when performing time analyses. The growth in the number of articles is, after all, not only the result of increased research activity, but also the effect of more extensive monitoring of scientific production - in this case, the greater number of journals included in WoS.

The greater openness of WoS to journals outside the global research centre is clearly visible in Central and Eastern Europe. In the years 2000-2006 published titles in CEE constituted barely more than 3% of all journals published in the entire EU27. As a consequence of the database's expansion, in the years 2007-2009 this percentage rose by around 3.5% and in subsequent years (2010-2013) maintained a level of around 7.5% (see Figure 1, left). This growth in the number of journals affected all the CEE countries analysed (see Figure 1, right), although the scale varied due partly to the size of the country as well as to the number of journals from a given countries that were listed in WoS before the 'global expansion' of 2007-2009. When we compare figures for 2000 and 2013, the greatest growth in journal numbers in WoS was noted in Romania and Estonia – where the number of titles increased 10-fold. To a large extent, this results from the low starting point (i.e. a low number of journals from these countries in 2000). A spectacular 5-fold increase was recorded in Slovenia. In the case of Poland and Bulgaria, growth was just over 3-fold. Meanwhile, Poland is the clear leader in the group of

countries analysed regarding absolute numbers of titles – one in three journals from the EU10 in the database is a journal published in Poland. Hungary, Lithuania and the Czech Republic doubled the number of journals in WoS and Slovakia increased its share by 40%. Meanwhile, Latvia is a very unusual case, as in 2013 only two journals from this country were present in WoS, and its entire growth is attributed to the addition of just one title in 2007. Another specific case is Lithuania: in the years 2000-2001, not a single journal published in this country appeared in WoS while, as a result of the expansion, as many as 29 titles had been included by the end of the decade.

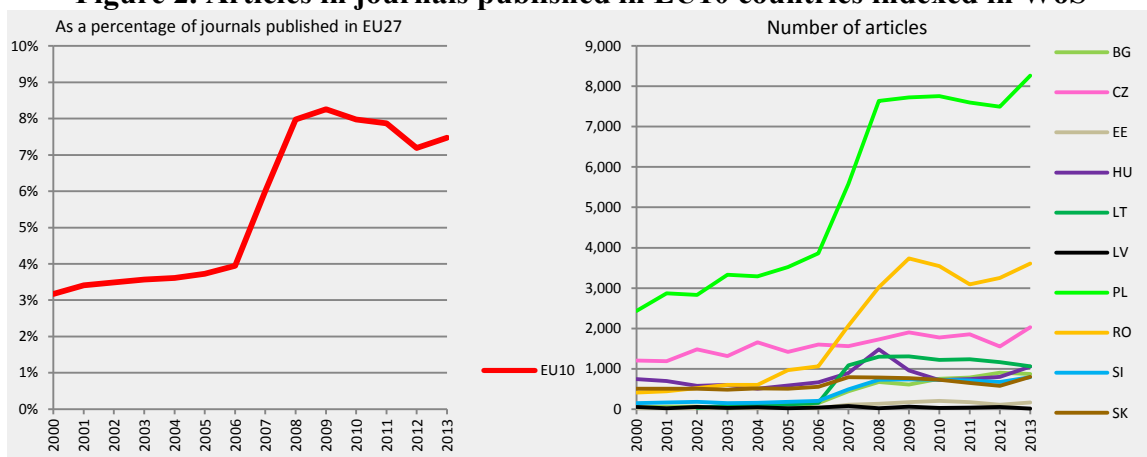
Analysis of the number of articles appearing in journals published in EU10 countries and included in WoS results in a similar picture to that given by analysis of the number of journal titles. Equally, in this case, there is a clear leap in the years 2007-2008 (Figure 2).

Figure 1. Articles published in EU10 countries and indexed in WoS



Source: own study based on data from the Web of Science.

Figure 2. Articles in journals published in EU10 countries indexed in WoS

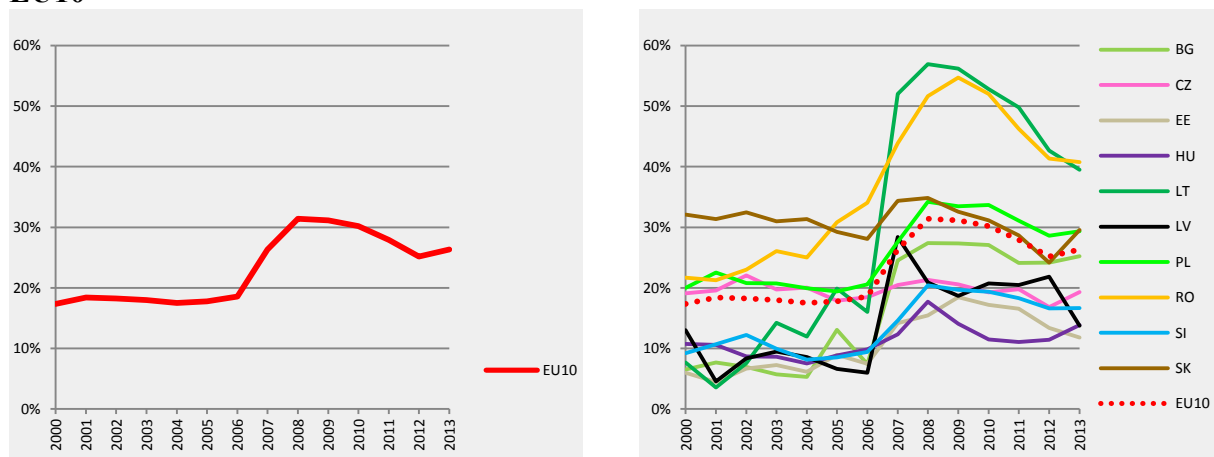


Source: own study based on data from the Web of Science.

One of the effects of including such a large number of national journals in WoS is the noticeable increase in the percentage of articles affiliated to EU10 countries and appearing in journals published by institutions in this part of Europe. In the years 2000-2006 these

constituted 17-18% of all articles affiliated to these countries and indexed in WoS. However, in the years 2008-2009 this percentage increased to 31%. Subsequent years saw a fall in numbers and in the years 2012 and 2013 only around one in four articles from the EU10 in WoS came from journals published in this region. Despite this drop, the figure is still higher than a decade earlier. The significance of national journals in the number of articles is very diverse in CEE. In the case of Lithuania and Romania, in the years directly following the expansion of WoS to include a large number of titles from these countries, more than half the articles in WoS came from journals published in the EU10 (which is almost equal to the number of journals from these countries – it is very rare that articles affiliated to the EU10 appear in journals published in other countries). However, the Czech Republic presents a completely different scenario. In this country, despite the number of journals in WoS doubling, the percentage of articles appearing in journals published in the EU in the years 2000-2013 (this also differs from national journals) remained at a level of 20% (see Figure 3).

Figure 3. Percentage of articles affiliated to EU10 appearing in journals published in EU10



Source: own study based on data from the Web of Science.

The difference in the share of articles appearing in journals published in countries in the region could testify to the differing levels of internationalisation of publishing activity. A high percentage of publications in journals of a particular country (region) can be interpreted as indicating a lower level of internationalisation in the science sector of this country. Meanwhile, a dominance of articles published in journals outside the country (region) in question indicates a greater presence in international research circles.

CONCLUSION

The aim of this article is to show that there is also a third factor, which influences the observable (but real?) convergence in term of research and development outcomes, in the form of publications in scientific journals. This factor is the wider inclusion of research articles from journals published in Central and Eastern European countries in worldwide bibliometric databases. To illustrate this, we used the Web of Science, which has for decades been the main reference source for international bibliometric comparisons. The evident impact made by inclusion, in 2007-2008, of numerous CEE journals in WoS on the values of the indicator analysed, directly supports the argument put forward in this article. The growth in the number of articles from individual countries in WoS thus not only testifies to the organic growth of the science sector in these countries but also results from decisions made by

the managing bodies of these commercial databases. Changes in the database are doubtless content-driven and are prompted, for example, by the desire to better reflect the state of world research. But we cannot reject other, non-content related motives, such as the wish to make the database more attractive to potential clients in countries with ‘developing’ science sectors and who have a great need to evaluate their research achievements (which in many countries goes hand in hand with reforms in higher education and science). However, irrespective of the reasons for expanding the database, it has led to better visibility in the European arena (Vanecek2014) and easier access to the research outcomes of CEE countries. And in the context of scientometric studies and international comparisons, we can say that the state of research in CEE countries is also being better measured.

REFERENCES

- EC (European Commission) (2014). *Innovation Union Scoreboard 2014*.
- Kosanovic, B., & Sipka, P. (2013). Output in WoS vs. representation in JCR of SEE Nations: does mother Thomson cherish all her children equally. In P. Šipka (ed.) *Journal Publishing in Developing, Transition and Emerging Countries: Proceedings of the 5th Belgrade International Open Access Conference 2012* (pp. 125-137). Belgrade: Centre for Evaluation in Education and Science.
- Testa, J. (2011). *The Globalisation of Web of Science: 2005-2010*. Thomson Reuters.
- Vanecek, J. (2014). The effect of performance-based research funding on output of R&D results in the Czech Republic. *Scientometrics*, 98(1), 657–681.



Does size matter? An investigation of how department size and other organizational variables influence on publication productivity and citation impact

Dag W. Aksnes,* Kristoffer Rørstad** and Fredrik N. Piro***

* Dag.W.Aksnes@nifu.no, ** Kristoffer.Rorstad@nifu.no, ***Fredrik.Piro@nifu.no

NIFU – Nordic Institute for Studies in Innovation, Research and Education, Økernveien 9, NO-0653 Oslo (Norway)

ABSTRACT

In this study, we investigate whether university department size is important in determining publication productivity and citation impact. Drawing on a unique dataset containing a variety of different variables at department levels, we are able to provide a richer picture of the research performance than what typically has been the case in many previous studies. In addition to analyzing the basic question of how size relates to scientific performance, we address whether the funding profile of the departments plays a role, whether the scientific performance is influenced by the composition of the academic personnel (in terms of gender, academic positions, recruiting personnel and the share of doctoral degree holders). The study shows that virtually no size effect can be identified and highly productive and highly cited units are found among both small, medium and large departments. For none of the organizational variables we are able to identify statistically significant relationships in respect to research performance at an overall level. We conclude that the productivity and citation differences at the level of departments cannot generally be explained by the selected variables for department size, funding structure and the composition of scientific personnel.

INTRODUCTION

In recent years, many countries have seen a strong emphasis and encouragement towards merger of research units – both within and across institutions (European Commission, 2009). Underlying this development are beliefs that larger departments are more cost-effective, reduces the administrative costs, and have advantages for both the study programs and research activities carried out. In addition to increased economies of scale and scope, rationales such as creating institutions that more effectively are capable of dealing with particular challenges typically are put forward as justifications for mergers in higher education (Goedegebuure, 2012).

There is also a widespread belief that the quality of the research suffers when the units are too small and that the number of researchers should be above a certain threshold (see e.g. Vehlo, 2006). The concept of “critical mass” is often used in this context. When applied in research policy, the expression alludes to an acceleration of the productivity or quality of a research unit above a certain size threshold.

The empirical justification for the size policy, however, does not seem to be particularly strong. A review of mergers in higher education in the early 1990s concluded that their

rationale often is based on questionable assumptions as to the expected outcomes in terms of increased economies of scale and scope (Goedegebuure, 1992). Similarly, a study performed by SPRU concluded that “there seems to be little if any convincing evidence to justify a government policy explicitly aimed at a further concentration of research resources on large departments or large universities in the UK on the grounds of superior economic efficiency” (von Tunzelmann et al., 2003).

Also when analyzing publication productivity specifically, there is little empirical evidence for the benefit of size. In a review of the literature almost twenty years ago, Johnston (1994) concluded that “research output is linearly related to size with no significant economies of scale apparent”. Later on, von Tunzelmann et al. (2003) concluded along the same lines. More recently, Evidence, in a commissioned report for the University Alliance (2011), found no evidence of critical mass in an analysis of the relationship between department size and various publication measures; both small and medium-sized research units tended to be at least as productive as large units. These results do not support the common assumption that department size in itself is beneficial for research productivity. Bonaccorsi & Daraio (2005) found even support for an opposite pattern where the most productive institutes in almost all areas were the small ones.

In this study of the Norwegian research system, we attempt to obtain further insights into the relationship between department size and scientific performance. Drawing on a unique dataset containing a variety of different variables at department levels, we are able to provide a richer picture of the research performance than what typically has been the case in many previous studies. In addition to analyzing the basic question of how size relates to scientific performance, we will address whether the funding profile of the departments plays a role, whether the scientific performance is influenced by the composition of the academic personnel (in terms of gender, academic positions, recruiting personnel and the share of doctoral degree holders). Most previous studies have analyzed the question focusing on publication productivity while there are fewer studies that have investigated research quality and size. Using data on the citation rate of the publications, we are able to assess both the publication productivity and scientific impact of the units. Finally, due to Norway’s good national research documentation system, all fields of learning may be included in our analyses, which is unusual in productivity studies.

More specifically, the following supplementary research questions and hypotheses are addressed:

The role of external funding

The institutions receive the majority of their funding through general government grants (i.e. internal funding). In the study, we will investigate whether the units receiving high proportions of external funding are more productive and have higher citation impact than the other units. The argument in favour of the hypothesis would be that those members of academic staff who apply for external funding have to document their past ability to publish their research. In the competition for research funding, the number of publications by applicants has become an important criterion for being worthy of future funding.

The qualifications of the academic staff

Generally, personnel holding doctoral degrees would be expected to be better qualified and prepared for an academic career than people lacking such qualifications. A previous study

based on Norwegian data also showed that academic staff holding a doctorate are more productive in terms of scientific and scholarly publishing than other staff (Kyvik & Aksnes, 2015). In the study, we will therefore assess whether there is a positive relationship between the proportion of the staff holding doctoral degrees and the units' academic performance in terms of productivity and citation impact.

The composition of the scientific staff

Several previous studies have shown that the average productivity of publications varies significantly across academic positions. For example, Rørstad and Aksnes (2015) showed that the publication rate of associate professors is generally 20-30 per cent below the one of the full professors, while the publication rate is lowest for PhD students. Based on such previous findings, one might assume that units with high proportions of full professors will have higher publication rates than the other units, while high proportions of PhD students will be negative associated with productivity.

The gender composition

There is strong evidence that female researchers tend to publish fewer publications than their male colleagues. This pattern has been found across many fields and nations (see e.g. van Arensbergen, van der Weijden & van den Besselaar, 2012). Less is known about whether similar gender differences can also be found in terms of citations. The few studies that actually have investigated this issue have not provided consistent results (Gonzalez-Brambila & Veloso, 2007) although a recent study of Norwegian researchers showed that females on average were cited slightly less than men (Aksnes et al., 2011). Based on such previous findings one might expect that departments with high rate of females fare less well when it comes to productivity and possibly citation impact.

The supplementary research issues described above are analysed at the level of departments. Many of the studies referred to have, however, been carried out at the level of individuals. It remains to be seen whether the patterns identified at an individual level also are evident at the aggregated department level.

DATA AND METHODOLOGY

Data on the research input are obtained using national R&D statistics. A variety of different variables at departments and institute levels are available through this statistic. In this study, the size of the departments and institutes is measured as number of work-years R&D. The latter numbers include time devoted to R&D, only. Thus, teaching and other activities are excluded. The figures are therefore suitable as a measure of the research efforts carried out and allow comparisons of units with different distributions of research and education.

The study is based on 210 units representing departments at universities and specialized university institutions in Norway. For the majority of the units analysed, we have three observations and the total number of observations underlying the analysis is 565.

Data on publication output are based on a bibliographic database called CRISStin, which is a common documentation system for all institutions in the higher education sector in Norway. CRISStin has a complete coverage of the scientific and scholarly publication output of the institutions.

In order to obtain an indicator that adjusts for different publication practices, we have calculated publication output as article equivalents. First, co-authored publications are fractionalised according to the number of authors. Based on this principle each department's share of a publication is calculated. Second, monographs are weighted as equal to five articles (in journals or books) in order to make the research efforts behind different types of publications comparable. The weighting of monographs corresponds to the principle applied in the Norwegian and Danish performance based funding model.

It should be noted that there is a time lag from the research is carried out until the research appear as published articles (usually one to two years or longer). In the study, we have applied a two-year time lag as a proxy. For most of the units, we have annual publication counts covering the period 2011-2013, while input data are from 2009-2011.

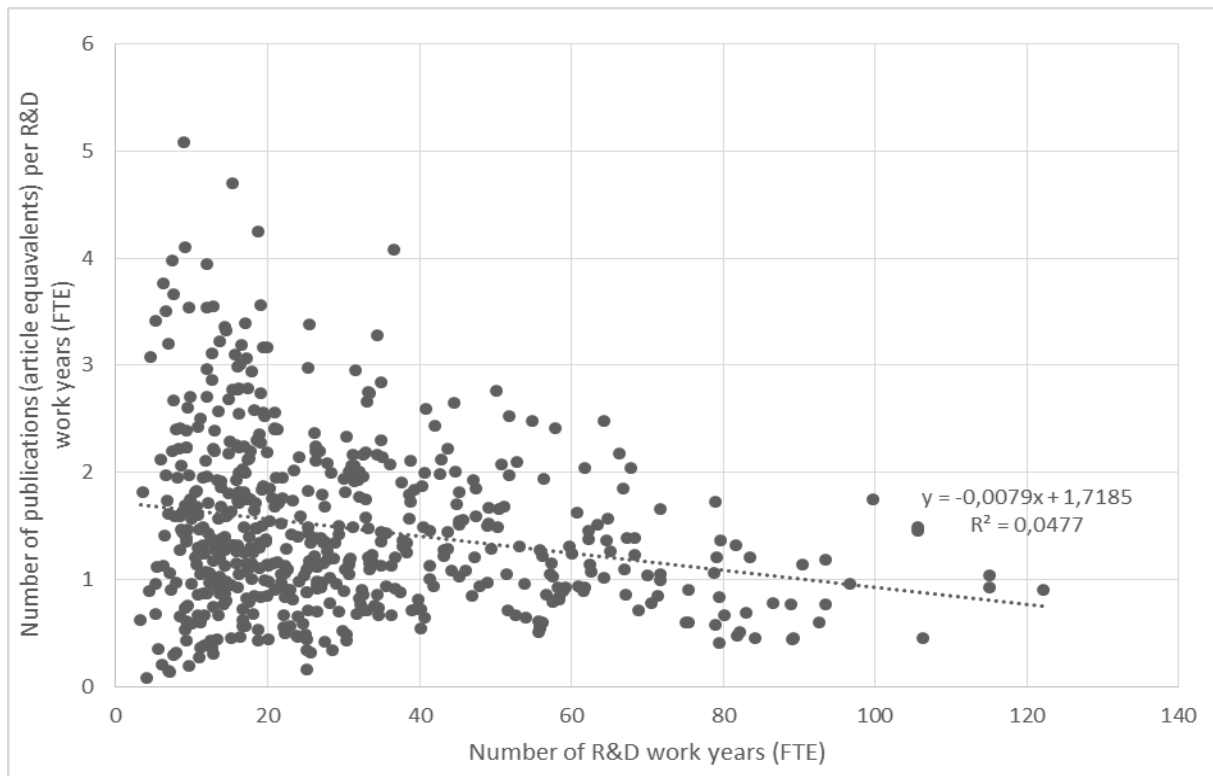
In addition to productivity measures, citation indicators have been calculated using data from Thomson Reuters' Web of Science (WoS) database. Many publications in CRIStin are not indexed in WoS. Therefore, this analysis is based on a more limited dataset. A threshold value of minimum 10 WoS articles annually was adopted in order to obtain reliable citation figures. Moreover, departments within the humanities are excluded, due to the publication and citation pattern of the field. The analysis is based on the articles published in the period 2011-2013 and citations obtained through 2014.

In the study, we have normalised the citation counts using the average citation rates of field and year in which the particular papers have been published. A citation indicator is subsequently calculated as the ratio of the citation rate of the articles to the average subfield citation rates. For example, an index value of 1.50 would mean that the articles are cited 50 per cent more frequently than the average.

RESULTS

In order to analyze whether larger departments have a higher relative productivity than smaller departments, we calculated the number of article equivalents per number of work years R&D (FTEs). The results are shown as a scatterplot (Figure 1) where the number of R&D work years (FTEs) is used as input variable. There are very large variations in the average productivity at department levels, particularly for the small departments. However, virtually no size effect can be identified and highly productive units are found among both small and medium sized departments. The linear regression line has a slightly negative slope with an R^2 value of only 0.05. There is no indication that a critical mass or a threshold value is present. Among the units with highest productivity, we find several small departments. This is probably due to the presence of one or a few highly prolific researchers, who influence significantly on the average of their small departments.

Figure 1. The relationship between department size (number of work years R&D (FTEs)) and productivity (number of article equivalents per work year R&D (FTEs)). (N=565).



In the study, we have used article equivalents to adjust for different publication practices. However, a previous study, partly based on identical data material, has shown that the indicator is not neutral across disciplines/domains (Piro, Aksnes, & Rørstad, 2013). On average, a researcher in the social sciences and humanities obtains significantly higher productivity rates than researchers in other fields, using this formula. We have therefore performed an analysis, taking into account the domain of the departments.

When plotting the size of the departments against the productive level (cf. Figure 1), we get a negative slope for four out five domains and only for the humanities the productivity increases with department size. However, the correlation is very weak with R^2 values in the range of 0.00-0.14. Thus, the conclusions above remain also when adding this variable to the analysis.

As a next step, we investigated how the composition of research personnel and the sources of funds were related to the publication output and department size. To be able to reveal any association between these variables, linear regressions were conducted (cf. Table 1). All the dependent variables are in the range zero to one, except the department size intervals (1-11).

Table 1. Regression summary for publications productivity (log of article equivalents per FTEs) and selected variables by domains. (N=565).

	Humanities	Social sciences	Natural sciences	Technology	Medical and health sciences	Total
Number of observations	91	218	98	87	71	565
R^2	<i>0.50</i>	<i>0.06</i>	<i>0.42</i>	<i>0.49</i>	<i>0.25</i>	<i>0.09</i>
General university funds	2.08*	0.03	0.10	1.09*	-0.79	0.49
Research Council of Norway funding	0.85	-0.14	-0.75	-0.34	-0.63	-0.33
Professors	-1.73	0.95	1.17	0.72	-0.08	0.10
PhD-students	1.54	-0.32	-0.03	1.67*	-0.31	0.58
Men	-0.05	-0.75	1.48*	0.99*	-0.40	0.03
Doctoral degree holders	3.15*	-0.12	-1.81	0.77	0.78	0.58
Department size (interval)	-0.14	-0.06	-0.10	-0.24	-0.04	-0.10

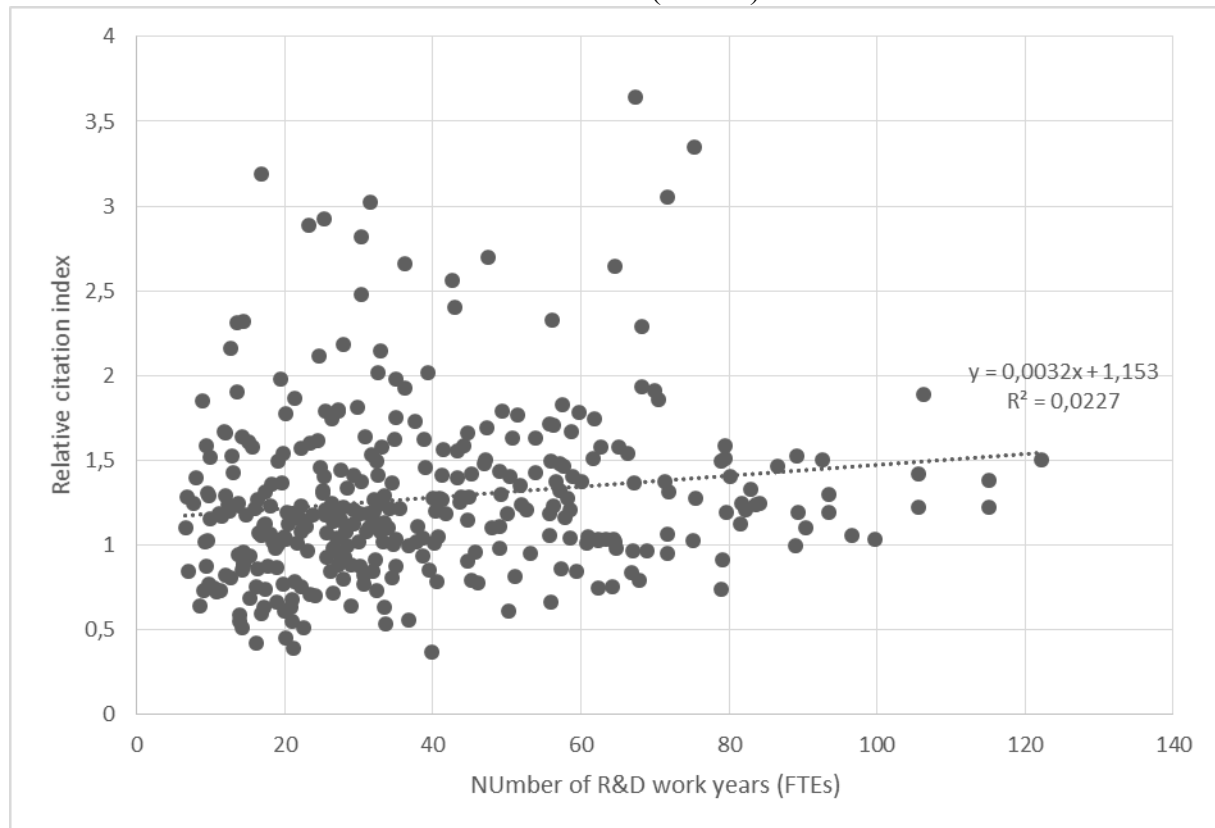
*) Statistically significant, 95% conf. interval.

The regression results show that the publication productivity can only partly be explained by our selected variables for department size, funding structure and the composition of scientific personnel. For the humanities, the natural sciences and technology, we obtain a fairly good correlation between publication productivity and the dependent variables, with R^2 values in the range of 0.42 to 0.50. Thus, these variables explain about half of the variance in publication productivity. For medical and health sciences, the variables explain about one fourth ($R^2 = 0.25$), while for the social sciences the R^2 value is 0.06, only. All domains, except the social sciences, have higher values than the total (0.09). The reason is that almost all the included variables have both a positive and a negative association with the publication productivity across domains.

For none of the variables we are able to identify statistically significant relationships in respect to publication productivity at an overall level. However, at the level of domains, a few statistically significant relationships have been identified.

As a next step, we analysed how the performance of the departments in terms of citation rates relates to their size. The relative citations index versus departments size (number of R&D work years (FTE)) is shown in Figure 2. The majority of the units are cited above the world average (1.00). However, also with this indicator, there is no systematic pattern. The trend line shows a slight increase in citation rate by size, but the correlation coefficient is very low with an R^2 value of only 0.02. Moreover, there are clearly no signs of any breakpoints where larger departments are more cited. Also when analysing the relationship at the level of domains there is no distinct pattern and the citation indexes do not systematically vary by size.

Figure 2. The relationship between department size (number of work years R&D (FTEs)) and citation index. (N=338).



A regression analysis also shows that the citation differences at the level of departments cannot generally be explained by the selected variables for department size, funding structure and the composition of scientific personnel. None of the independent variables have a statistically significant association with the citation index, this also holds for the relationship between department size and citations.

DISCUSSION

We are not able to identify any benefits of larger departments when it comes to research performance measured through bibliometric indicators. To the contrary, the study shows that a) there are no systematic productivity or citation differences between small, medium and large departments, and, b) there is no evidence of a critical mass or a threshold level. As described in the introduction, similar findings have also been found in many previous studies.

There may be several reasons for this apparent tension between the empirical results and the presumed benefits of larger departments. Prolific research groups may be found within both large and small departments. Possibly, the pros of having larger departments only are influential in some of the cases. In addition, both small and large departments may have their advantages and disadvantages.

In the study, we are focusing on the formal organizational level: the department. Larger departments, in particular, typically have sub-departments, which may operate quite independent of each other. When using the department as the only variable, internal differences in the organizational structure are concealed. Moreover, it is the research groups

that are the functional units of science. Previous studies indicate that the group is more important than the department in explaining research productivity (von Tunzelmann et al., 2003).

We have also included other departmental variables in the analyses. Overall, the included variables explain 9 per cent of the variance in publication productivity and 10 percent of the variance in citation index. For none of the variables we are able to identify statistically significant relationships in respect to research performance at an overall level. We conclude that the productivity and citation differences at the level of departments cannot generally be explained by the selected variables for department size, funding structure and the composition of scientific personnel.

Therefore, the hypotheses stated in the introduction cannot be sustained by the empirical findings of the study. This is surprising as several of the variables have been shown to be influential at the level of individuals (cf. Introduction). Apparently, patterns present at the level of individuals are concealed when aggregated departments are analyzed. This means that at this level other factors are more important for explaining the variance in publication productivity and citation rates. Thus, one has to look at other aspects of the organizational structure in order to explain these differences.

CONCLUSIONS

Knowledge concerning factors influencing on the scientific performance of scientists and research organizations is important in research policy and management. Based on such knowledge one may be able to create better research conditions and design effective organizations to increase productivity and fostering high quality research. Our study, does not give support to the widespread policy assumption that small departments in this respect are unfavorable compared to larger departments. Although there may be arguments in favour of larger departments along other dimensions, the lack of empirical support when it comes to research performance is an important finding to bring forward in discussions about the organization of the higher education systems.

REFERENCES

- Aksnes, D.W., Rorstad, K., Piro, F. & Sivertsen, G. (2011). Are Female Researchers Less Cited? A large-scale study of Norwegian scientists. *Journal of the American Society for Information Science and Technology*, 62(4), 628-636.
- van Arensbergen, P., van der Weijden, I. & van den Besselaar, P. (2012). Gender differences in scientific productivity: a persisting phenomenon? *Scientometrics*, 93(3), 857-868.
- Bonaccorsi, A. & Daraio, C. (2005). Exploring size and agglomeration effects on public research productivity. *Scientometrics*, 63(1), 87-120.
- European Commission (2009). *The Question of R&D Specialization: Perspectives and Policy Implications*. Luxemburg: European Commission.
- Goedegebuure, L. (2012). *Mergers and More: The changing tertiary education landscape in the 21st century*. HEIK Working Paper Series. 2012/1. Oslo: University of Oslo.

- Goedegebuure, L.C.J. (1992). *Mergers in Higher Education; A comparative perspective*. Utrecht: Lemma.
- Gonzalez-Brambila, C., & Veloso, F. M. (2007). The determinants of research output and impact: A study of Mexican researchers. *Research Policy*, 36(7), 1035–1051.
- Johnston, R. (1994). Effects of resource concentration on research performance. *Higher Education*, 28(1), 25-37.
- Kyvik, S. & Aksnes, Dag W. (2015). Explaining the increase in publication productivity among academic staff: a generational perspective. *Studies in Higher Education*, 40(8), 1438-1453.
- Piro, F, Aksnes, D. W. & Rørstad, K. (2013). A macro Analysis of Productivity differences across fields: Challenges in the Measurement of Scientific Publishing. *Journal of the American Society for Information Science and Technology*, 64(2), 307-320.
- von Tunzelmann, N., Ranga, M., Martin, B. & Guena, A. (2003). *The Effects of Size on Research Performance: A SPRU Review*. Brighton: University of Sussex.
- University Alliance (2011). *Funding research excellence: research group size, critical mass & performance*. London: University Alliance.



Do usage and scientific collaboration associate with citation impact?

Pei-Shan Chi ^{*} and Wolfgang Glänzel ^{**}

^{*} peishan.chi@kuleuven.be; wolfgang.glanzel@kuleuven.be
ECOOM and Dept MSI, KU Leuven, Naamsestraat 61, Leuven 3000, (Belgium)

^{**} glanzw@iif.hu
Library of the Hungarian Academy of Sciences, Dept. Science Policy and Scientometrics, Arany János u. 1,
Budapest, 1051 (Hungary)

ABSTRACT

In this study usage counts and times cited from Web of Science Core Collection (WoS) were collected for each article published in 2013 with Belgian, Israeli and Iranian addresses. We investigate the relations among three indicators related to citation impact, usage counts co-authorship, respectively. In addition, we apply the method of Characteristic Scores and Scales (CSS) to analyse the distributions of citations and usage counts. The results show that citations and usage counts in WoS correlate to each other significantly, especially in the social sciences. However, the increase of the number of co-authors does not increase usage counts or citations significantly. Furthermore, the stability of CSS-class distributions proves the availability of CSS in characterising both usage and citation distributions.

INTRODUCTION

According to Brody, Harnad and Carr (2006), a reading-citing cycle of scholarly publications from the moment an article is accepted for publication until it is published, read or cited may range from 3 months to 1–2 years or even longer. The usage impact being measured earlier in the reading-citing cycle is significant and may be predictive for the later stage of the cycle, i.e. the citation impact which can be only measured after the publications of those citing articles of a given article. Among the 39 scientific impact measures tested by Bollen et al. (2009), usage-based measures are even stronger indicators of scientific prestige than many citation measures.

Beyond this reading-citing cycle, scientific collaboration may increase the quality of the research. Thus, it may influence the above measures of the quality of scholarly publications. The number of authors is often used as a measure of the scientific collaboration. For example, Peters and van Raan (1994) detected a general correlation between citation counts and number of authors. In this study, we investigate the relation among citations, usage counts and the number of authors per paper.

METHODOLOGY

We collect usage counts and number of citations from WoS for three countries with similar publication output as pars pro toto examples. Relevant data were extracted for each article published in 2013 from two developed country (Belgium and Israel) and a developing country (Iran). 28,746 papers with at least one Belgian address each were downloaded on October 16, 2015, while 30,906 Iranian papers and 19,837 Israeli papers were collected on March 1, 2016

and March 15, 2016, respectively. The usage count reflects the number of times a WoS user clicking links to the full-length article at the publisher's website or saving the article for use in a bibliographic management tool. It is counted since from the beginning of usage counting on 1 February 2013 till the date of data download.

All items extracted from the WoS have been assigned to 17 major fields according to the *modified* Leuven-Budapest classification system (Glänzel & Chi, 2015). In this study, three major fields, Mathematics, Social Sciences II (economic, political & legal studies), Clinical and Experimental Medicine II (non-internal medicine specialties) are selected to analyse the relations between usage, citation impact and scientific collaboration (see Table 1). All the samples are further analysed for the correlation coefficients among citation counts, usage counts and the number of co-authors per paper, and the distributions of citations and usage counts by *Characteristic Scores and Scales* (CSS) (Glänzel & Schubert, 1988).

Table 1. Statistics of sample data sets.

	Mathematics			Social Sciences II			Clinical & Experimental Medicine II		
	BEL	ISR	IRN	BEL	ISR	IRN	BEL	ISR	IRN
Total numbers of papers	908	925	1,808	1,071	593	132	5,814	5,041	3,220
Mean citation rate	2.53	2.20	3.12	2.39	2.70	3.91	4.90	5.04	2.39
Mean usage rate	12.06	6.20	10.55	20.56	19.32	26.02	9.52	8.01	7.71
Co-authors per paper	2.96	2.40	2.56	3.02	2.31	2.79	7.47	6.31	4.83

RESULTS

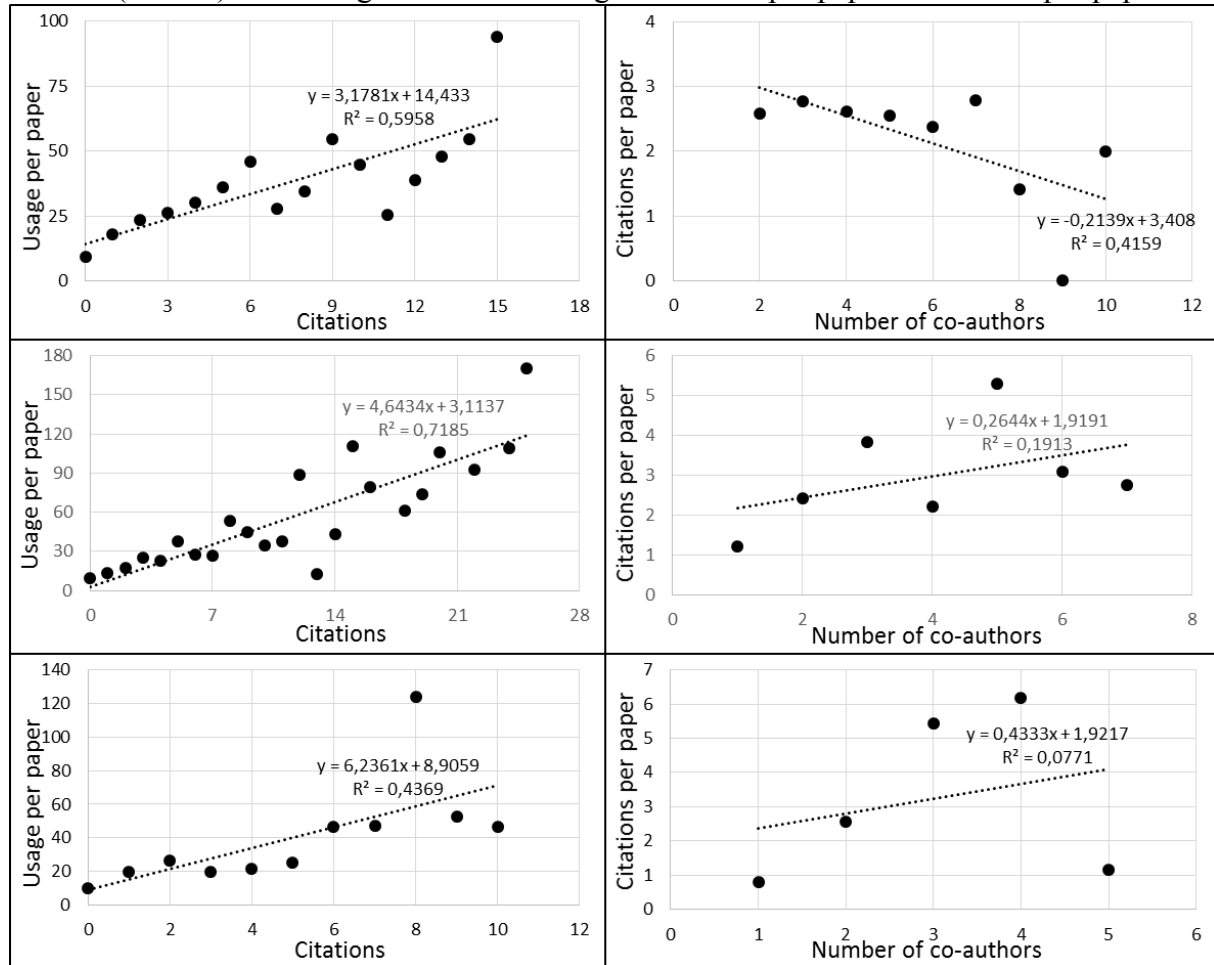
Table 2 shows that the correlation between citations and usage counts are medium but much higher than other two groups, especially in the social sciences. This finding is related to the fact shown in Table 1 that the usage amounts comparing to citations in the social sciences are larger than the other field. In contrast, WoS papers in clinical medicine are not used as much as other fields probably because of another popular medicine data source, PubMed.

Table 2. Person correlation coefficients of publications in Belgium, Israel and Iran.

		Cites vs. Usage	Co-authors vs. Usage	Co-authors vs. Cites
Mathematics	BEL	0.3475690***	0.0766077*	0.0435901
	ISR	0.3228998***	0.1421174***	0.1911563***
	IRN	0.3901064***	0.1957826***	0.1575937***
Social Sciences II	BEL	0.4758493***	0.0928735**	0.1105789***
	ISR	0.4011023***	0.1767232***	0.3401602***
	IRN	0.6184363***	0.1061911	0.1763957*
Clinical & Experimental Medicine II	BEL	0.4100875***	0.1140624***	0.1717247***
	ISR	0.5622304***	0.1080933***	0.1676701***
	IRN	0.4514596***	0.0486313**	0.2399343***

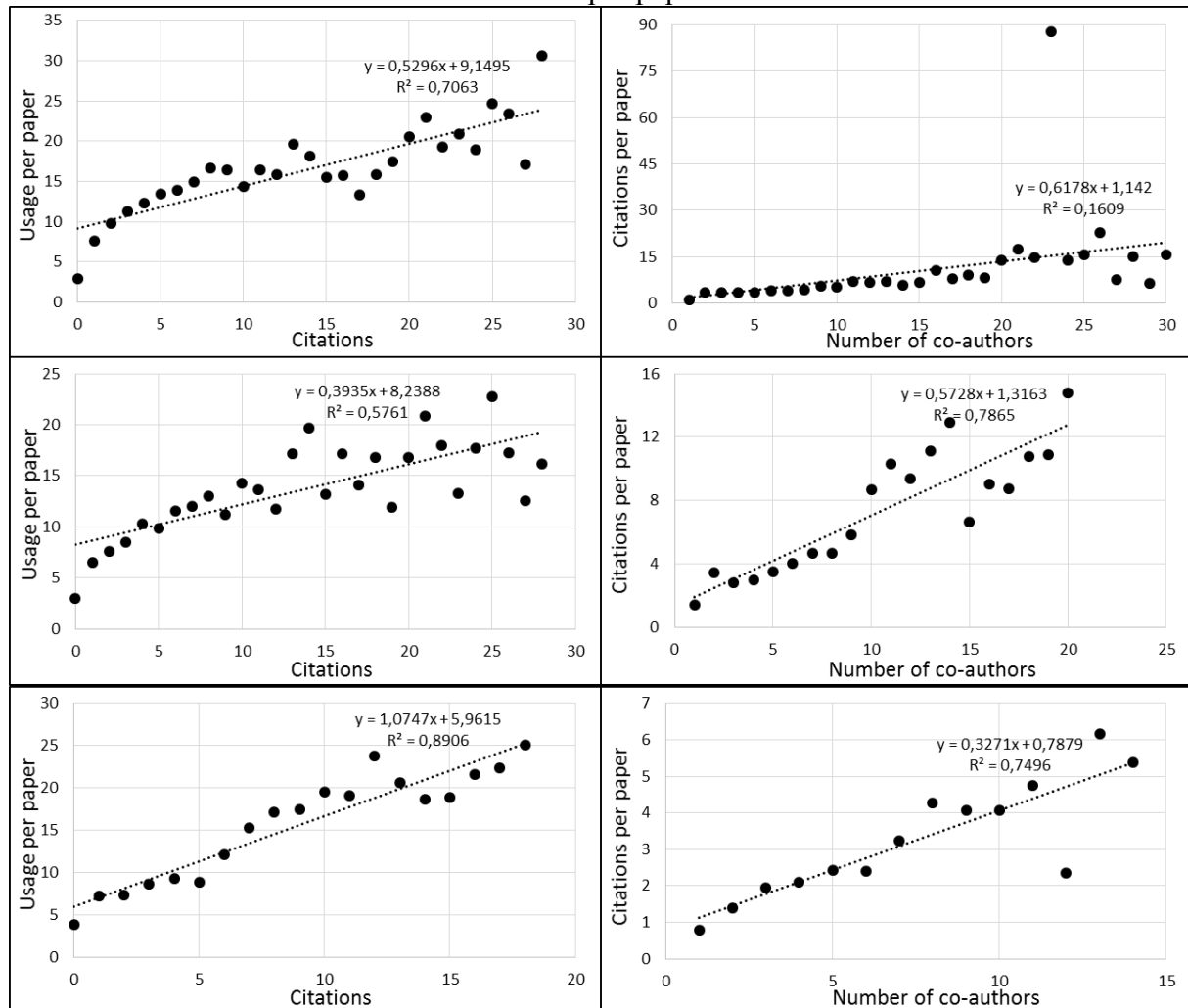
* p -value < .05, ** p -value < .01, *** p -value < .001

Figure 1: Scatter plots of publications in Social Sciences II. Belgium (top), Israel (centre) and Iran (bottom). Left: usage vs. citations. Right: citations per paper vs. authors per paper.



Providing a closer look at the stronger relations between citations and other indicators, Figures 1 and 2 illustrate the scatter plots of major publications in the social sciences and clinical medicine based on the regression of usage on citations and citations on number of co-authors. In order to avoid distortions caused by outliers we have limited the number of citations and co-authors, respectively, to an appropriate scale. In the social sciences, Figure 1 shows that all the three countries have positive correlations between citations and usages. However, they have distinct correlations between the number of co-authors and citations, displaying an negative relation in Belgian publications versus positive ones in Israeli and Iranian publications. In clinical medicine, Figure 2 shows that Israel and Iran have stronger links in both correlation between usage and citations, and correlation between citations and numbers of authors than Belgium.

Figure 2: Scatter plots of publications in Clinical & Experimental Medicine II. Belgium (top), Israel (centre) and Iran (bottom). Left: usage vs. citations. Right: citations per paper vs. authors per paper.



This is the first time CSS is applied to usage data. For citation data, the distribution of papers of CSS-classes roughly obeys the 70%–21%–6.5%–2.5% rule (from the lowest to the highest class – see Glänzel et al., 2014). The stability of this property shown in Table 3 proves the availability of CSS in characterising usage distributions as well. Apart from the distribution over classes, also the scores themselves reveal interesting aspects. On one hand, we observe distinctly different patterns in citations and usage, where, according to the expectations, usage has, in general, higher scores than citations and, on the other hand, we see some variations among the three countries though, but the similarities within citations and usage in these fields are somewhat surprising. For example, publications in the social sciences have the highest usage numbers compared to other fields, especially in Iran. It indicates the role of WoS in the social sciences as search tool more than target journal set to publish. Compared to other countries, Iran has relatively lower highly-cited scores in mathematics and clinical medicine, but relatively higher highly-used scores in the social sciences.

Table 3. Characteristic scores and CSS-class shares of publications (C – citations, U – usage counts) for three countries: Belgium (top), Israel (centre) and Iran (bottom).

Class	Mathematics				Social Sciences II				Clinical & Experimental Medicine II			
	Score(C)	%(C)	Score(U)	%(U)	Score(C)	%(C)	Score(U)	%(U)	Score(C)	%(C)	Score(U)	%(U)
1	0.0	68.4	0.0	75.1	0.0	70.4	0.0	64.9	0.0	69.0	0.0	67.9
2	2.5	22.1	12.1	17.7	2.4	20.1	20.6	22.4	4.9	22.9	9.5	20.6
3	6.5	6.7	33.8	5.0	6.6	6.7	43.7	8.2	13.4	6.3	20.5	7.2
4	12.4	2.8	69.2	2.2	12.2	2.8	68.8	4.5	29.6	1.8	33.4	4.3
Class	Score(C)	%(C)	Score(U)	%(U)	Score(C)	%(C)	Score(U)	%(U)	Score(C)	%(C)	Score(U)	%(U)
1	0.0	73.5	0.0	73.4	0.0	69.6	0.0	70.2	0.0	75.6	0.0	70.2
2	2.2	19.4	6.2	19.5	2.7	22.1	19.3	19.9	5.0	18.3	8.0	20.0
3	6.5	4.9	16.5	4.3	7.4	6.1	44.7	6.4	16.6	4.8	19.0	6.9
4	13.1	2.3	33.7	2.8	15.8	2.2	77.4	3.5	38.4	1.3	32.6	2.9
Class	Score(C)	%(C)	Score(U)	%(U)	Score(C)	%(C)	Score(U)	%(U)	Score(C)	%(C)	Score(U)	%(U)
1	0.0	73.0	0.0	67.2	0.0	70.5	0.0	68.9	0.0	70.4	0.0	68.2
2	3.1	19.4	10.6	22.5	3.9	21.2	26.0	20.5	2.4	19.9	7.7	23.4
3	9.1	4.9	24.4	7.3	11.0	5.3	57.2	7.6	6.7	6.8	17.9	6.0
4	17.9	2.8	42.9	3.0	22.8	3.0	100.1	3.0	12.3	2.8	36.7	2.4

CONCLUSIONS

We found a significant but not strong correlation between citations and usage counts in WoS. However, similarly to earlier observations concerning the correlation between downloads and citations, no causality in one particular direction should be assumed (Glänzel & Heffeer, 2014). The analysis of the relation between ‘usage’ and download might be one of the future research tasks.

On the other hand, the increase of number of authors does not increase usage counts or citations as significantly as how usage counts correlate with citations. It is especially notable in the case of Belgian publications in the social sciences. In turn, Israeli publications in the social sciences have the most similar associations among three indicators. The three countries in our sample set show different patterns of three relations. In short, Belgium generally has weaker correlations among these three indicators than other two countries.

The application of CSS was proved to work in the usage counts as well as citation counts, keeping the stability of class distributions between citation and usage. Additionally, distinctly different patterns in citations and usage are observed, but the similarities within citations and usage in these fields are somewhat unexpected. Social sciences has the most distinct patterns with much higher usage than citations compared to other fields, revealing the function of WoS in this field as search tool instead of target journal set to publish. The three examples substantiate that, in general, there is no clear relationship between WoS usage and citation counts; the sometimes contradicting relationship between number of co-authors and citation impact, however, surprises.

REFERENCES

- Bollen, J., Van de Sompel, H., Hagberg, A. & Chute, R. (2009). A Principal Component Analysis of 39 Scientific Impact Measures. *PLoS ONE*, 4(6), e6022.
- Brody, T., Harnad, S. & Carr, L. (2006). Earlier Web Usage Statistics as Predictors of Later Citation Impact. *JASIST*, 57(8), 1060–1072.

Glänzel, W. & Schubert, A. (1988). Characteristic scores and scales in assessing citation impact. *Journal of Information Science*, 14(2), 123–127.

Glänzel, W., Thijs, Debackere, K. (2014). The application of citation-based performance classes to the disciplinary and multidisciplinary assessment in national comparison and institutional research assessment. *Scientometrics*, 101 (2), 939–952.

Glänzel, W. & Heefffer, S. (2014). *Cross-national preferences and similarities in downloads and citations of scientific articles: A pilot study*. In: E. Noyons (Ed.), Proceedings of the STI Conference 2014, Leiden, 207–215.

Glänzel, W. & Chi, P.S. (2015). *A comparative study of citation distributions of books and journal articles in three fields and implications for indicator building*. In: Proceedings of the 20th International Conference on Science and Technology Indicators, Lugano, 55–58.

Peters, H.P.F. & van Raan, A.F.J. (1994). On determinants of citation scores: a case study in chemical engineering, *JASIST*, 45(1), 39–49.

An approach for the condensed presentation of intuitive citation impact metrics which remain reliable with very few publications

David Campbell*, Chantale Tippett*, Grégoire Côté*, Guillaume Roberge* and Éric Archambault*

*david.campbell@science-metrix.com; chantale.tippett@science-metrix.com; gregoire.cote@science-metrix.com; guillaume.roberge@science-metrix.com; eric.archambault@science-metrix.com
Science-Metrix, 1335 Avenue du Mont-Royal Est, Montréal, H2J 1Y6 (Canada)

Abstract

An approach for presenting citation data in a condensed and intuitive manner which will allow for their reliable interpretation by policy analysts even in cases where the number of peer-reviewed publications produced by a given entity remains small is presented. The approach is described using country level data in Agronomy & Agriculture (2004–2013), an area of specialisation for many developing countries with a small output size. Four citation impact metrics, and a synthesis graph that we call the distributional micro-charts of relative citation counts, are considered in building our “preferred” presentation layout. These metrics include two indicators that have long been used by Science-Metrix in its bibliometric reports, the Average of Relative Citations (ARC) and the percentage of publications in the 10% most cited publications in the database (HCP), as well as two newer metrics, the Median of Relative Citations (MRC) and the Relative Integration Score (RIS). The findings reveal that the proposed approach combining the MRC and HCP with the distributional micro-charts effectively allows to better qualify the citation impact of entities in terms of central location, density of the upper citation tail and overall distribution than Science-Metrix former approach based on the ARC and HCP. This is especially true of cases with small population sizes where a strong presence of outliers (denoted by strong HCP scores) can have a significant effect on the central location of the citation data when estimated with an average.

Keywords

Citation impact metrics; Average of Relative Citations; Median of Relative Citations, Relative Integration Score; Highly Cited Publications; central location/position; research excellence, distributional micro-charts of relative citation counts

Introduction

Historically, Science-Metrix has used two complementary citation impact metrics in assessing the scientific impact of various entities (e.g. countries, provinces/states and organisations) (for an example, see Campbell et al., 2013). The Average of Relative Citations (ARC)—which is similar to the Mean Normalized Citation Score (MNCS) recently introduced by the Center for Science and Technological Studies (CWTS) (Waltman et al., 2011)—was traditionally used to infer how a given entity was generally performing. The percentage of publications in the 10% most cited publications in the database (i.e. Highly Cited Publications [HCP]) was used to assess the scientific “excellence” of the research performed by a given entity (Bornmann, Leydesdorff & Wang, 2013).¹

¹ These two indicators were sometimes completed with an *indirect* impact metric based on the impact factor of the publication venues (i.e. the Average of Relative Impact Factors) instead of the actual publications.

However, it is now widely recognised that scientific impact indicators relying on the mean of the papers' citation scores can be highly influenced by extreme values (denoted by strong HCP scores) due to the strong skewness observed in citation distributions (Albarran, Ortuño Ortin & Ruiz-Castillo, 2011; Campbell, 2011; Bornmann, Leydesdorff & Wang, 2013). Indeed, citation distributions are usually characterised by a large number of publications receiving zero citations and a small number obtaining the vast majority (Seglen, 1992). It results that a single measure of centrality may not be adequate to capture the upper and lower ends of the distribution.

This is why Science-Metrix traditionally presented the ARC, which is based on the mean, alongside the HCP, which adequately captures the upper end of the distribution. Unfortunately, since the ARC can be highly affected by the upper tail of the distribution, it does not always adequately reflect the central location in the citation data and can correlate strongly with the HCP. This issue is compounded in instances where the unit of evaluation in question possesses a small number of publications (e.g. small countries, large countries in very small research topics, research organisations or researchers). In fact, it appears, based on experimental data, that the effect of outliers on the rank of entities based on the ARC can be pronounced for entities with as much as 10,000 publications (Campbell, 2011).

As a result, isolates and/or peripheral countries within international co-publication networks were most often left aside in Science-Metrix bibliometric reports as the indicator toolset which was in use did not adequately capture their scientific impact; most studies focused on the largest producers. Yet, it can be of high interest to study the scientific performance of smaller producers such as low/middle income economies in areas of high strategic importance to their development. One such area in which developing nations have a very small number of publications while being highly specialised is the subfield of Agronomy & Agriculture.

Thus, the present study aims to develop an approach for presenting citation data in a **condensed** and **intuitive** manner which will allow for their reliable interpretation by policy analysts even in cases where the number of peer-reviewed publications produced by a given entity remains small; the analysis of country level data in Agronomy & Agriculture for the 2004–2013 period is used as a test case. We first review some of the most recently introduced alternatives to mean-based citation impact metrics. Subsequently, the methods used in computing two indicators that have long been used by Science-Metrix in its bibliometric reports, the Average of Relative Citations (ARC) and the percentage of publications in the 10% most cited publications in the database (HCP), as well as two newer metrics, the Median of Relative Citations (MRC) and the Relative Integration Score (RIS) are presented. The methods used in producing the distributional micro-charts of the relative citation scores of an entity are also presented. These micro-charts, which can be displayed within a traditional table alongside other indicators, are introduced to enable the intuitive comparison (visually) of the citation impact of entities accounting for the entire distribution of their citation scores. The analytical approach used in setting-up our “preferred” presentation layout for the analysis of citation impact is then presented by comparing the ranking of countries in Agronomy & Agriculture based on the above four indicators and the distributional micro-charts. A short discussion of the potential applicability of the newly developed presentation layout and of upcoming work concludes the paper.

Methods

Literature review

As with all methodologies, other methodologies that have been developed in order to address the issues of skewed citation distributions, such as Foster, Greer and Thorbecke's (FGT) (1984) family of low- and high-impact measures or percentile-based approaches, or the P100 metric of Bornmann, Leydesdorff, & Wang (2013), have drawbacks. For instance, while methodologically robust and providing a wealth of information on the high- and low tails of the citation distribution, the visual representation of the FGT family of measures is complex (with three measures to characterise both the FGT high- and low-impact families; six metrics in total) and thus may not be appropriate for evaluations conducted outside of a research setting. Indeed, in the private sector setting, a central aim is to support a sound methodological approach with clear and concise visuals that are easily interpretable to a wide audience. Additionally, while percentile-based approaches are now viewed as the most robust alternative to mean-based normalisation methods in computing citation impact metrics (see Leiden Manifesto [Hicks et al., 2015]), they often lead to numbers (e.g. average of the [percentiles] ranks) that are not highly communicative of the actual gaps in the citation counts of entities at the most central location (i.e. the 50th percentile or median), a clear drawback for disseminating the information to a wide audience. Indeed, as formulated by Bornmann, Leydesdorff, & Wang (2013), only the order, but not the scale, is preserved in the citation counts of papers using percentile-based approaches. Additionally, since they report a single number, that is the average of the (percentiles) ranks, they still hide an important portion of the variability in the underlying distribution of the citation scores across entities; for instance, two entities could share the same average rank while having underlying distributions that differ.

In this study, a central aim is to support a sound methodological approach with clear and concise visuals that are easily interpretable to a wide audience. To achieve this, two percentile-based tools (one metric and one visual) are introduced in an attempt to improve the interpretative value of percentile-based metrics as well as to allow for the condensed and intuitive display of the performance of entities along the full distribution of citation scores. The introduced metric is the Median of the Relative Citations (MRC) which allows for the formulation of highly communicative statements of the following form: "The median paper of entity x (e.g. Philippines) is cited 50% more often than the median world paper" (the indicator is subfield, year and document type normalised). The second goal is achieved by producing the distributional micro-charts of the relative citation counts which is also used in computing the Relative Integration Score (RIS); another percentile-based metric which accounts for the entire distribution but which is hard to interpret. These micro-charts, which can be displayed within a traditional table alongside other indicators, are introduced to enable the intuitive comparison (visually) of the citation impact of entities accounting for the entire distribution of their citation scores.

Data sources

Only a portion of the work performed on the whole of the selected databases² within the Web of Science (WoSTM; Thomson Reuters) is presented in this paper, namely for the Agriculture and Agronomy subfield for the 2004–2013 period. Only articles (and notes which no longer exit in WoS) and reviews have been considered. These two document types are collectively referred to as "papers" or "publications" in the remainder of this paper.

² Science Citation Index Expanded (SCI-Expanded) database; Social Sciences Citation Index (SSCI); Arts & Humanities Citation Index (AHCI)

The subfields used in normalising the indicators presented here are those found in Science-Metrix's journal-based classification of scientific research (Archambault, Caruso & Beauchesne, 2011). Agronomy & Agriculture is one such subfield in Science-Metrix's classification.

The following sub-sections present each of the impact indicators presented in this papers starting with: the Average of Relative Citations (ARC) and the Median of Relative Citations (MRC). The distributional micro-chart of relative citation scores, and its associated metrics (i.e. the percentage of publications in the 10% most cited publications in the database [HCP] and the Relative Integration Score [RIS]) is then presented and. Finally, three other indicators which are used in analysing the results are briefly presented (due to space constraints): the number of papers, the activity index, and the collaboration index.

Average of Relative Citations (ARC)

The Average of Relative Citations (ARC) is an indicator of the scientific impact of papers produced by a given entity (e.g., a country, an institution) relative to the world average (i.e., the expected number of citations). The number of citations received by each publication is counted for the year in which it was published and all subsequent years (i.e. citation windows of variable length). To account for different citation patterns across scientific subfields (e.g., there are more citations in biomedical research than in mathematics) and document types (reviews include more references and are more cited than articles), as well as to account for differences in the age of publications (i.e., older papers have accumulated citations over a longer period), each publication's citation count is divided by the **average** citation count of all publications (in the WoS; sometimes the term "world" is used in reference to the whole of the database) of the corresponding document type that were published the same year in the same subfield. In this way, one arrives at a relative citation count (RC). The ARC of a given entity is the **average** of the RCs of the papers belonging to it. An ARC value above 1 means that a given entity is cited more frequently than the world average, while a value below 1 means that its publications receive on average fewer citations than the world average. A statement of the following form can be made for a country (e.g. Philippines) with an ARC of 1.45: "On average, the papers from the Philippines are cited 45% more frequently than the average world paper".

Median of Relative Citations (MRC)

The computation of the MRC is very similar to the computation of the ARC as detailed above, except that the averages in the ARC's definition are replaced by the medians in the MRC's definition. The MRC is an indicator of the central position of the citation impact of an entity's papers (e.g., a country, an institution) relative to the world (i.e., the expected number of citations at the 50th percentile). A MRC value above 1 means that a given entity's median paper is cited more frequently than the world median paper, while a value below 1 means that its median publication receives fewer citations than the world median paper. A statement of the following form can be made for a country (e.g. Philippines) with a MRC of 1.50: "The median paper from Philippines is cited 50% more frequently than the median world paper".

Distributional micro-charts of the relative citation scores

The first step in producing this distributional micro-chart consists in subdividing the world's papers (i.e. all the papers in the database) into ten subgroups each representing 10% of the total number of papers (i.e. in 10 deciles), with the papers sorted in ascending order of their number of received citations (from the least [1st decile] to the most [10th decile] cited papers).

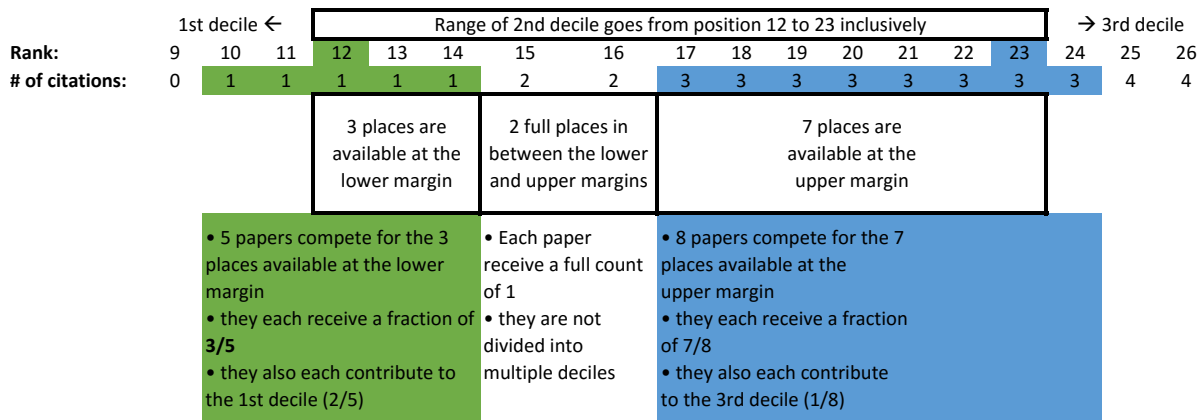
To allow for the subsequent aggregation/comparison of data from different subfields, document types and years, this segmentation is performed independently for each subfield/year/document type combination, thus creating the relative citation scores (within each combination, we work with the raw citation scores in subdividing the world's papers in ten equal subgroups).

Of course, because of tied citation scores, some papers will have to be fractionated across multiple deciles. Nevertheless, all papers with the same number of citations will be distributed across deciles in the same manner with the same fraction. Although Bornmann, Leydesdorff, & Wang (2013) argued that fractioning the papers should be avoided to enable the applicability of statistical tests, we disagree with them. For instance, it is possible to apply a z-test for two independent proportions to compare the shares of the scientific production of two entities which fall within the 10% most cited publications using fractionated paper counts. Indeed, the units of analysis in this case are not the papers themselves, but rather the shares in the top 10%—which can be computed as the sum of an entity's paper fractions falling in the top 10% over the entity's total number of papers—and the sample sizes—which can be computed as an entity's total number of papers. That being said, if an analysis aims to directly compare individual papers, then fractioning should indeed be avoided.

The approach used in subdividing the world's papers into 10 citation deciles is illustrated for a fictitious subfield, year and document type combination including a total of 113 papers in the world (i.e. in the WoS). In this hypothetical case, each decile should therefore include 11.3 papers (i.e. $113/10$). The splitting process takes place in four steps and is illustrated in Figure 1 below:

- **Step 1:** Order the papers according to their number of received citations (distinct rank even for ties; see Figure 1 below).
- **Step 2:** Identify the lower and upper margin of each decile (this example is for the 2nd decile, but the same approach applies to all 10 deciles).
 - Lower margin of 2nd decile = upper margin of 1st decile plus one; i.e. $\text{ROUND}(0.1 * 113, 0) + 1 = 12$.
 - Upper margin of 2nd decile = $\text{ROUND}(\text{decile} * 113, 0) = \text{ROUND}(0.2 * 113, 0) = 23$.
- **Step 3:** Determine the fraction of a given paper which contribute to the 2nd decile (see explanation in the Figure 1 below).
- **Step 4:** Adjust the fraction attributed to each paper accounting for the number of places available in each decile (i.e. 11.3) relative to the actual number of places in a specific decile based on its computed margins (see above explanations on computing the margins; in the current example, there are 12 places in the 2nd decile):
 - Multiply each paper fraction by a constant;
 - Constant = number of available places per decile divided by the number of places in decile = $11.3/12$.

Figure 1: Subdividing the world’s papers in a given subfield, year and document type into 10 subgroups each representing 10% of the total number of papers (i.e. in 10 deciles) from the least (1st decile) to the most (10th decile) cited papers.



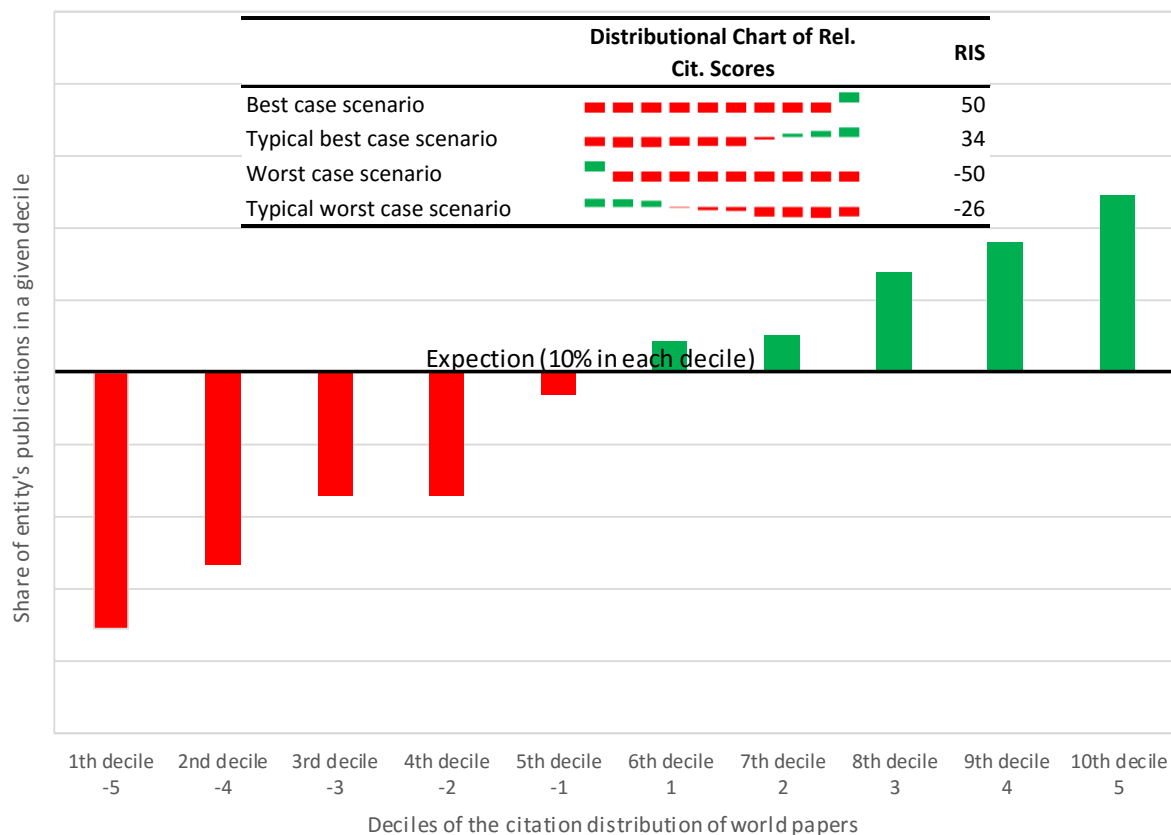
Note: Some deciles can be identical if there are lots of ties. This is most often the case for the lowest deciles due to the high occurrence of papers without any citation. In these cases, the papers are still fractionated into multiple deciles according to the above procedure.

Once this important step has been performed, it is possible to examine how the papers of a given entity are distributed across citation deciles. If the papers of an entity were drawn randomly from the world’s distribution, then we would expect to observe 10% of its papers in each decile (i.e. the expected proportion). Because countries (and other entities) diverge from one another in their citation impact, characterising the departure from expectation across all deciles allows one to appreciate whether an entity generally performs better or worse than the world in general accounting for the entire distribution of the relative citation scores.

This can be visualised by constructing a distributional chart of the relative citation scores of an entity. What is actually displayed in this chart is the ratio of the observed to the expected proportions of papers in each decile. The proportion of an entity’s papers in a given decile is equal to the sum of its paper fractions in this decile divided by the sum of its paper fractions across all deciles. To ensure comparability across entity, the scale of the micro-charts has been standardised between -1 and 1 with 0 representing the world level. This is achieved by taking the hyperbolic tangent of the natural logarithm of the ratio of the observed to the expected proportions in each decile. The resulting display is shown in Figure 2 below.

The expected share in each decile is 10% and is marked by the absence of a bar in the graph (the bold gridline). Scores below expectation are coloured in red, while scores above expectations are coloured in green. The best performances are marked by the presence of red bars to the left (more observations than expected in the low citation deciles) and of green bars to the right (more observations than expected in the high citation deciles). The opposite pattern marks the worst case scenarios. Typically, a positive slope is indicative of good overall performance while a negative slope is indicative of a mitigated (“bad”) performance; a slope of 0 usually indicate a performance near expectations across all deciles (except for extreme cases; i.e. best or worst case scenarios, see Figure 2).

Figure 2: Distributional chart of the relative citation scores (year, subfield and document type normalised) of various entities (e.g. countries) and the computation of the Relative Integration Score (RIS)



Although the computations which are required to produce such charts are quite complex, the algorithm we developed in SQL language to interact directly with our internal implementation of the WoS ran in just a few hours for the entire database (for all subfields, document types and years). Additionally, the resulting charts is simple, intuitive and allows comparing a large number of entities, alongside other indicators, in a single table when drawn has micro-charts using Excel sparklines tool (see Table 1 in the results section).

Relative Integration Score (RIS)

Using the above distributional chart, it is possible to derive a percentile-based indicator (i.e. the RIS) which account for the performance of entities over the entire citation distribution. Below the axis label for each decile in Figure 2 is the weight used in computing the RIS. The ratio of observed to expected share in each decile (non-standardized) is multiplied by the corresponding decile weight and these products are then summed to obtain the RIS. The RIS ranges from -50 (worst case scenario) to 50 (best case scenario) with 0 representing the world level.

Highly cited publications (HCP)

The percentage of publications of an entity that falls in the 10% most cited publications in the database is the score (in its non-standardised form) shown in figure 2 for the 10th decile. It corresponds to the sum of an entity's paper fractions that fall in the 10th decile divided by the sum of the corresponding entity's paper fraction across all deciles. A score above 10% marks a performance above the world level while the opposite is true of a score below 10%.

Number of papers

This indicator shows the number of publications for a given country, using full counting.

Activity index (AI; or specialisation index)

The AI, or specialization index (SI), indicates how much emphasis a given country puts on one subfield, relative to the global average of effort exerted in that subfield. For instance, if 20% of a country's publications are in Agronomy & Agriculture, but at the global level only 15% of papers are in this subfield, then the country is said to be specialized with an AI of 1.33 (i.e. $0.20/0.15$). A SI above the world level of 1 indicates relative specialisation, whereas a score below 1 indicate relative de-specialisation.

Collaboration index

The CI is the ratio of the number of observed international co-publications to the number expected according to the non-linear relationship between the number of papers and the number of international co-publications of countries. A CI above 1 denotes higher-than-expected rates of international collaboration for a given country, a CI below 1 denotes the opposite, and a CI close to 1 denotes a rate of collaboration near expectation (the world level).

Results & Discussion

A Spearman's rank correlation matrix between the computed impact indicators was first computed for countries in Agronomy & Agriculture (Table 1). From this matrix, one can readily see that impact measures based on the mean (i.e. the ARC) are much more strongly correlated with HCP than those based on the median (i.e. the MRC), highlighting the strong influence of the upper tail of citation distributions on the former indicator. It is thus quite obvious that the MRC is a better measure of central location in citation distributions than the ARC is.

Also notable is the fact that the strongest correlation is observed between the MRC and the RIS ($R^2 = 0.93$) which fulfils all of the desirable properties of indicators describing entire citation distributions (symmetry, replication invariance, continuity, focus, monotonicity and normalisation, Albarran, Ortuño Ortin & Ruiz-Castillo, 2011); note that in our case, we have ten scores (one per decile) merged into one instead of two (one for low and high impact papers). Thus, the MRC, which is based on the median, appears to reflect quite decently the general performance of countries when accounting for the entire distribution as is done with the RIS. Also, because both measures correlate similarly poorly with the HCP, the MRC and the RIS appear, compared to the ARC, as better complementary measures to the HCP; they are less redundant than the ARC providing information on the general performance, in terms of citation impact, of an entity rather than focussing on "excellence" (i.e. on highly cited publications). However, because the MRC is much simpler and intuitive than the RIS—and because it can be converted into a textual finding which can easily be interpreted by a wide audience (every policy analyst is familiar with the median)³—our "preferred" presentation layout would consist of pairing the MRC with the HCP in analysing the scientific impact of entities to, respectively, assess their general performance in scientific impact as well as to investigate their level of research "excellence". These measures would be combined with the distributional micro-charts of the relative citation scores to add nuance in the interpretation of the findings (see below discussion of the results presented in Table 2).

³ A MRC of 1.50 for the Philippines means that "the median paper from the Philippines is cited 50% more frequently than the median world paper".

Still, the correlation of the ARC with the MRC and RIS remains quite high (Table 1). This is not surprising either since with the law of large numbers, the various measures will converge at some point; for many of the large producers, the ranks will remain somewhat stable regardless of the selected impact metrics. The main added value of the MRC (and RIS) is to help detect when the presence of high impact papers is sufficient to alter the average as a good measure of central tendency. This is especially important when analysing the impact of small producers such as the low/middle income economies in Agronomy & Agriculture as presented in Table 2.

Table 1. Spearman's rank correlation matrix between various citation impact metrics in Agronomy & Agriculture (2004–2013)

	ARC	MRC	HCP	RIS
ARC	1.00			
MRC	0.79	1.00		
HCP	0.81	0.51	1.00	
RIS	0.86	0.93	0.57	1.00

Note: ARC = Average of Relative Citations; MRC = Median of Relative Citations; HCP = Highly Cited Publications; RIS = Relative Integration Score

The main outliers in the relationship between country ranks based the ARC and MRC are highlighted with a tick border in Table 2. They include Malawi with 83 publications and Kenya with 589 papers. Comparing the ranks based on the ARC and the MRC, Malawi lost 31 places in the ranking and Kenya lost 20 places. This drop reflects the fact that the ARC of both countries, as a measure of central tendency, was inflated by the presence of more papers than expected in the top citation deciles (i.e. the 7th, 8th and 10th deciles for Malawi and the 9th and 10th decile for Kenya; see the distributional micro-charts). In fact, the rank of Kenya based on the ARC is similar to its rank based on the HCP; this is less the case for Malawi as two other deciles (i.e. 7th and 8th), beyond the 10th one (i.e. HCP), contribute to this effect (only the 9th and 10th for Kenya). Also interesting to note is the fact that the rank of both countries based on the MRC and the RIS are similar; they are identical for Kenya. Also worthy of mention is the fact that countries with similar ARC scores actually have better overall performance than Malawi (i.e. Israel, Costa Rica, Syria and the US) and Kenya (i.e. Sri Lanka and Morocco) when looking at the MRC and the RIS. This can be appreciated by looking at their distributional micro-charts. In the case of Israel, Costa Rica, Syria, the US, Sri Lanka and Morocco, there is a stronger concentration of green bars in the higher citation deciles (the five deciles to the right) and of red bars in the lower citation deciles (i.e. the five deciles to the left) compared to Malawi and Kenya. This indicates a better overall performance.

The proposed approach also enables us to present data for countries with very few publications without fear of erroneously concluding that they globally perform really well in citation impact when this would only be due to a few strong outliers in the upper tail of the distribution. For instance, one can see that Panama, with only 47 publications, stands out as the country with some of the strongest scores across the board (5th in ARC, 4th in MRC, 6th in HCP and 2nd in RIS). The strength of this performance, in spite of the small production size, is exemplified by the distributional micro-charts which shows that Panama's papers are highly concentrated in the top 3 deciles; 79% of its papers are in the 30% most cited papers in the world. Given that it is highly specialised in this subfield (SI of 2.79), it remains highly pertinent to develop the required toolset to appropriately assess its scientific performance in

spite of its very small production. Since it is unlikely that such a small production could have raised as much impact, due to the Matthew Effect in science, it is also of interest to perform additional analysis to shed light on this unique case. A likely explanation lies in international co-authorships. For instance, Panama collaborated more than expected with foreign partners (CI of 1.40), perhaps in the context of large consortium which will normally raise very large citation impacts. The same holds true for many developing nations in Table 2 (e.g. Zimbabwe).

We conclude by stating that our “preferred” visual layout holds the promise to ease the communication of reliable data on the scientific impact of small entities down to the organisational and researcher levels. Note that we are also currently performing additional analyses, using simulations, to measure the critical value at which point the population size (i.e. number of papers) of an entity is sufficiently large for the central limit theorem to apply as a function of the ARC; it is anticipated that the critical population size at which point the ARC converges towards the MRC will increase as the ARC increases.

Table 2. Scientific performance of countries in Agronomy & Agriculture based on various citation impact metrics, the activity index and the collaboration index (2004–2013)

Country	Papers	AI	CI	ARC		MRC		HCP		RIS		Distributional Chart of Rel. Cit. Score
				Score	Rk	Score	Rk	Score	Rk	Score	Rk	
Panama	47	2.79	1.40	2.81	5	4.67	4	46.0%	6	33.9	2	
Switzerland	881	1.62	1.50	2.13	6	3.00	7	28.4%	8	25.2	6	
Zimbabwe	167	3.72	1.44	2.04	7	2.20	19	26.2%	11	15.7	20	
Austria	547	1.80	1.36	2.02	8	2.75	8	23.1%	16	19.4	13	
Lao	55	4.49	1.57	2.02	8	2.00	22	20.9%	22	18.4	15	
Sweden	1,112	1.45	1.25	1.98	10	2.50	12	25.7%	12	22.1	9	
Ireland	408	1.98	1.08	1.94	11	2.47	15	23.9%	15	21.6	10	
Denmark	1,257	2.11	1.18	1.87	12	2.63	11	24.2%	14	23.0	8	
Netherlands	2,050	1.65	1.56	1.86	14	2.11	21	21.6%	18	18.1	18	
United Kingdom	3,604	1.26	1.54	1.82	15	2.25	18	22.7%	17	18.2	16	
France	3,107	1.50	1.46	1.74	17	2.50	12	21.4%	20	21.6	10	
Belgium	1,199	1.96	1.41	1.73	18	2.13	20	21.3%	21	18.2	16	
Madagascar	45	3.77	1.56	1.72	20	2.40	16	28.2%	9	23.8	7	
Germany	4,801	1.48	1.30	1.71	21	2.00	22	21.6%	18	15.5	21	
Italy	2,239	0.91	1.05	1.65	22	2.00	22	19.3%	25	16.2	19	
New Zealand	1,374	2.55	0.96	1.50	24	1.76	31	17.1%	31	13.0	28	
Portugal	527	1.86	1.18	1.49	25	2.00	22	18.9%	26	14.1	23	
Finland	831	1.32	0.66	1.47	26	2.00	22	17.9%	28	13.6	26	
Australia	5,267	2.15	0.99	1.46	27	1.67	32	15.5%	35	12.8	29	
Spain	4,148	1.26	0.95	1.45	28	2.00	22	17.8%	29	14.2	22	
Philippines	594	3.04	1.30	1.45	28	1.50	37	18.6%	27	6.5	51	
Israel	603	1.62	0.94	1.41	31	1.83	30	15.8%	34	14.0	24	
Malawi	83	3.30	1.38	1.38	32	1.08	63	13.2%	48	5.8	53	
Costa Rica	72	2.85	1.57	1.36	34	1.63	33	19.5%	24	12.0	30	
Syria	252	4.51	1.46	1.36	34	1.50	37	13.5%	47	9.0	44	
United States	19,319	1.06	0.91	1.36	34	1.50	37	14.7%	37	10.4	36	
Singapore	35	1.50	1.19	1.33	38	1.63	33	16.9%	32	11.4	31	
China	8,354	1.56	1.14	1.33	38	1.56	35	14.9%	36	9.6	41	
Uzbekistan	77	4.19	1.50	1.31	40	1.50	37	13.9%	44	9.4	43	
Mali	83	3.51	1.69	1.30	42	1.50	37	10.9%	59	9.5	42	
Sri Lanka	198	3.94	1.48	1.29	43	1.39	50	14.0%	43	5.1	57	
Kenya	589	4.51	1.63	1.28	44	1.00	64	14.4%	38	2.8	64	
Morocco	121	3.07	1.45	1.26	46	1.50	37	13.6%	46	10.5	35	
Norway	466	1.64	1.14	1.25	47	1.33	52	12.3%	52	8.2	46	
Canada	4,694	1.79	0.95	1.21	48	1.44	49	12.4%	51	8.3	45	
Zambia	30	2.45	0.81	1.21	48	1.13	60	10.0%	64	2.4	65	
Senegal	122	5.40	1.67	1.20	50	1.53	36	9.2%	73	9.8	40	
Mexico	842	2.24	1.33	1.14	51	1.25	55	12.5%	50	5.1	57	
Colombia	262	2.36	1.51	1.13	52	1.00	64	10.2%	63	0.3	76	
Argentina	1,006	2.52	0.80	1.12	54	1.50	37	10.0%	64	7.8	48	
Indonesia	243	2.89	1.70	1.12	54	1.00	64	14.4%	38	2.1	67	
...												
Ecuador	33	2.69	1.49	0.98	73	1.00	64	12.8%	49	1.0	71	
Ethiopia	336	6.32	1.40	0.97	76	1.00	64	9.7%	67	0.7	73	
Greece	735	1.63	0.61	0.93	81	1.00	64	8.9%	76	0.8	72	
...												
Nigeria	1,153	3.66	0.62	0.32	143	0.20	142	1.7%	124	-25.8	147	
India	6,558	2.18	0.30	0.48	125	0.20	142	4.6%	106	-20.2	136	
Bulgaria	322	1.62	0.41	0.31	145	0.13	148	3.1%	116	-25.6	145	

Note: AI = Activity Index; CI = Collaboration Index; ARC = Average of Relative Citations; MRC = Median of Relative Citations; HCP = Highly Cited Publications; RIS = Relative Integration Score. Only countries with an average production of at least 3 papers per year over the study period are considered (total of 30 papers). Not all these countries are shown for space reason. Countries are sorted in descending order on the basis of their ARC.

References

- Albarran, P., Ortuño Ortin, I., & Ruiz-Castillo, J. (2011). The measurement of low- and high-impact in citation distributions: Technical results. *Journal of Informetrics*, 5(1), 48–63. doi:10.1016/j.joi.2010.08.002
- Archambault É., Caruso J., and Beauchesne O. (2011). Towards a Multilingual, Comprehensive and Open Scientific Journal Ontology, in Noyons, B., Ngulube, P. and Leta, J. *Proceedings of the 13th International Conference of the International Society for Scientometrics and Informetrics (ISSI)*, Durban, South Africa, pp 66-77.
- Bornmann, L., Leydesdorff, L., & Wang, J. (2013). Which percentile-based approach should be preferred for calculating normalized citation impact values? An empirical comparison of five approaches including a newly developed citation-rank approach (P100). *Journal of Informetrics*, 7(4), 933–944.
- Campbell, D. (2011). Dimensionality Reduction in the Multi-Criteria Analysis of Scientific Performance Does research performance correlate with environment-related outcomes of countries? *4S Annual Meeting*, Cleveland, Ohio, http://www.science-matrix.com/pdf/SM_Campbell_4S_Composite_Index_Scientific_Performance.pdf.
- Campbell, D., Lefebvre, C., Picard-Aitken, M., Côté, G., Ventimiglia, A., Roberge, G. & Archambault, E. (2013). *Country and Regional Scientific Production Profiles*. Report prepared for DG-RTD by Science-Matrix, ISBN: 978-92-79-33735-2, <https://rio.jrc.ec.europa.eu/en/file/7183/download?token=skuVMnRo>.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S. & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, 520, 429-431.
- Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science*, 43(9), 628–638.
- Waltman, L., van Eck, N.J., van Leeuwen, T.N., Visser, M.S. & van Raan, A.F.J. (2011). Towards a New Crown Indicator: Some Theoretical Considerations. *Journal of Informetrics*, 5, 37-47.



A comparison of average-based, percentile rank, and other citation impact indicators

Pedro Albarrán* and Javier Ruiz-Castillo**

*albarran@ua.es

Departamento de Fundamentos del Análisis Económico, Universidad de Alicante

** jrc@eco.uc3m.es

Javier Ruiz-Castillo, Departamento de Economía, Universidad Carlos III,

ABSTRACT

The main aim of this paper is to defend the view that, in spite of the broad agreement in favor of the *MNCS* and the *percentile rank* indicators, there are two other citation indicators with desirable properties that the above indicators do not possess: (i) a member of the family of high-impact indicators introduced in Albarrán *et al.* (2011), and (ii) a new indicator, based in the work of Herrero & Villar (2013), which measures the *relative* performance of the different research units in terms of a series of tournaments in which each research unit is confronted with all others repeatedly. We compare indicators from the point of view of their *discriminatory power*, measured by the range and the coefficient of variation. Using a large dataset indexed by Thomson Reuters, we consider 40 countries that have published at least 10,000 articles in all sciences in 1998-2003. There are two main findings. First, the new indicator exhibits a greater discriminatory power than percentile rank indicators. Second, the high-impact indicator exhibits the greatest discriminatory power.

Introduction

There is a general agreement among Scientometrics' practitioners about taking into account the citations received in research evaluation exercises. Our starting point is the following impression: historically, there has been a broad agreement concerning which indicators we should use. We distinguish between two periods.

- Since the 1980s, there has been a general agreement favoring average-based indicators of two types: the *crown indicator* and, after an intense discussion in the 2007-2011 period, the *Mean Normalized Citation Score* (MNCS hereafter). See *inter alia* Lunberg (2007), and Waltman *et al.* (2011).

- Since 2010, a group of highly qualified professional leaders and influential institutions have shown their current preferences for an alternative to average-based indicators that Bornmann & Marx (2012) call the *percentile rank approach*. (See also the *I3* indicator in Leydesdorff *et al.*, 2011, and Rousseau, 2012). In practice, given the skewness of citation distributions, authors give more importance to percentiles or percentile rank classes at the upper tail of the reference citation distribution. In this paper, we focus on the *Topk* indicators, defined as the percentage of an institution's scientific output included into the set formed by the $k\%$ most cited papers in a given scientific field. We refer to indicators that partition the reference citation distribution into two categories as *dichotomous* indicators.

This broad agreement manifests itself in the indicators used in the two more important reference rankings for Scientometrics practitioners: the *Leiden Ranking* (LR hereafter) for universities, and the *SCImago Institutions Ranking* for research institutions. Both use the MNCS and the *Top 10%* indicator. In addition, the 2015 edition of the LR use the *Topk* indicators for $k\% = 50\%, 1\%$.

The main aim of this paper is to defend the view that, in spite of the broad agreement in favor of the MNCS and the *Topk* indicators, there are two other citation indicators with desirable properties that the above indicators do not possess. First, Albarrán *et al.* (2011) introduced the notion of high-impact indicators defined over the publications with citations above a *critical citation line* (CCL hereafter). In this paper, we focus on a member of the family of FGT high-impact indicators whose properties are inherited from a class of economic poverty indicators introduced by Foster, Greer, & Thorbecke (1984). Second, in this paper we introduce a new indicator, based in the work of Herrero & Villar (2013), which measures the *relative* performance of the different research units in terms of a series of tournaments in which each research unit is confronted with all others repeatedly.

Any comparison between alternative indicators should involve their properties, as well as the empirical differences they give rise in applications. As far as the empirical criterion, in this paper we compare indicators from the point of view of their *discriminatory power*, measured by the range and the coefficient of variation (CV hereafter).

Citation impact indicators

Let N_{ij} be the number of articles of unit i in field j , and let c_{ij} be the ordered *citation distribution of unit i in field j* . Let c_j denote the ordered *citation distribution of field j* , where $N_j = \sum_i N_{ij}$ is the total number of articles in field j . Consider the first three types of indicators that will be used in this paper.

1. Let μ_{ij} and μ_j be the mean citation of distributions c_{ij} and c_j , respectively. The *Relative Citation Rate*, RCR , is defined as

$$RCR_{ij} = \mu_{ij}/\mu_j. \quad (1)$$

For field j as a whole, $RCR_j = \mu_j/\mu_j = 1$, and $\sum_i (N_{ij}/N_j)RCR_{ij} = RCR_j = 1$.

For dichotomous indicators, consider the partition of c_j into two categories: publications in the top k percentiles, and publications in the bottom $(100 - k)$ percentiles. Alternatively, fix the CCL, z_j^k , equal to the number of citations of the article in the $(1 - k)^{\text{th}}$ percentile of citation distribution c_j . For any article l in research unit i in field j , the *CCL normalized high-impact gap* is defined as

$$\gamma_{ijl} = \text{Max} [0, (c_{ijl} - z_j^k)/z_j^k]. \quad (2)$$

Note that $\gamma_{ijk} > 0$ only for high-impact articles with citations $c_{ijk} > z_j^k$. Consider the family of FGT high-impact indicators whose properties are inherited from a class of economic poverty indicators introduced by Foster, Greer, & Thorbecke (1984):

$$H_{ij}^\beta = [1/N_{ij}] \sum_l (\gamma_{ijl})^\beta, \beta \geq 0, \quad (3)$$

where β is a parameter identifying the members of the family. For the entire field j as a whole, $H_j^\beta = [1/N_j] \sum_i \sum_l (\gamma_{ijl})^\beta$.

2. It turns out that the first member of this family for $\beta = 0$, referred to as Top^k , is precisely the Top^k indicator:

$$H_{ij}^0 = Top_{ij}^k = n_{ij}^k/N_{ij}, \quad (4)$$

where n_{ij}^k is the number of high-impact articles published by unit i . For field j as a whole, $Top_j^k = n_j^k/N_j = 0.k$, where $n_j^k = \sum_i n_{ij}^k$.

3. The second member of this family for $\beta = 1$ is the average of the normalized high-impact gaps, referred to as ANG^k :

$$H_{ij}^1 = ANG_{ij}^k = (1/N_{ij})(\sum_l \gamma_{ijl}). \quad (5)$$

For the entire field j as a whole, $ANG_j = (1/N_j)[\sum_i \sum_l \gamma_{ijl}]$.

4. Finally, the worth approach is presented in four steps. (i) For any pair of research units, u and v , let p_{uv}^k be the probability that a publication in unit u is in a higher position than a publication in unit v . With only two categories, the probability p_{uv}^k can be easily computed as follows:

$$p_{uv}^k = T_u^k (1 - T_v^k).$$

Similarly, p_{vu}^k denotes the probability that a publication in unit v is in a higher position than a publication in unit u , where $p_{vu}^k = T_v^k (1 - T_u^k)$. We say that unit u *dominates unit v in a pair-wise comparison* when $p_{uv}^k > p_{vu}^k$ or, in other terms, when $T_u^k > T_v^k$. Thus, the ranking provided by this criterion coincides with that given by the *Topk*. In the dichotomous case, this is a very reasonable condition. However, the indicator values differ. The reason is that the *Topk* indicator evaluates research units in an absolute sense, while in the worth approach the evaluation is relative. Let us see how this works.

(ii) The relative value of two research units u and v is given by the quotient

$$\frac{p_{uv}^k}{p_{vu}^k} = \frac{T_u^k (1 - T_v^k)}{T_v^k (1 - T_u^k)} = \frac{T_u^k - T_u^k T_v^k}{T_v^k - T_v^k T_u^k}.$$

Now, in order to extend this relative evaluation criterion to a more general setting, involving any finite number of research units, we need some additional elaboration since now we have to take into account all pairwise dominations simultaneously. The *relative advantage of unit u with respect unit v* , r_{uv}^k , is defined as follows:

$$r_{uv}^k = p_{uv}^k / \sum_{s \neq u} p_{sv}^k.$$

That is, r_{uv}^k is the ratio between the probability of unit u dominating unit v in a pair-wise comparison, and the sum of the probabilities of unit u being dominated by any other unit.

(iii) A natural way of assigning a global evaluation to unit u , e_u , is by means of a weighted average of its relative advantages:

$$e_u = \sum_{v \neq u} \lambda_v r_{uv}^k,$$

where λ_v is a measure of the importance attached to unit v .

(iv) We are interested in the case where the importance attached to the different units derives, precisely, from the importance that the evaluation procedure yields. We refer to such a system of weights as the *worth vector* $\mathbf{Worth}^k = (Worth_{1, \dots, u, \dots, l}^k)$ where, for each k ,

$$Worth_u^k = \sum_{v \neq u} Worth_v^k r_{uv}^k.$$

Herrero & Villar (2013) establish that such a system of weights always exists and is unique once the scale has been chosen. We set the scale so that $\sum_u Worth_u^k = 1$. The intuitive rationale of the new indicator is that the worth of a unit is higher, all other things being equal, the higher the worth of the units it dominates.

(A numerical example with the four types of indicators is available on request)

Properties of the indicators

Table 1 gives a schematic view of indicator properties. Formal definitions and a discussion can be found in the references at the bottom of Table 1.

Table 1. Schematic view of properties with an indication of which of them, when applicable, are or are not satisfied by the four types of indicators

TYPE OF INDICATOR:	<i>RCR</i>	<i>Topk</i>	<i>ANGk</i>	<i>Worthk</i>
Scoring rule	Yes	Yes	Yes	No
Relative	No	No	No	Yes
PROPERTIES:				
A. Replication-invariance	Yes	Yes	Yes	Yes
B. Scale-invariance	Yes	Yes	Yes	Yes
C. Translation-invariance	No	Yes	No	Yes
D. Additive decomposability	Yes	Yes	Yes	Not applicable
E. Subgroup consistency	Yes	Yes	Yes	Not applicable
F. Independence	Yes	Yes	Yes	No

G. Strict monotonicity	Yes	No	Yes	No

References: Properties **A B C D E G**: Albarrán *et al.* (2011); Property **F**: Marchant (2009) and Waltman & van Eck (2009)

The following three comments are in order.

- In the first three cases, an indicator is a real valued function whose domain is the space of all citation distributions, and whose range is the set of non-negative real numbers. It can be shown that they are all scoring rules in the sense of Marchant (2009). Scoring rules provide an *independent* evaluation of each research unit, i.e., the value attached of each unit is always the same, irrespective of the remaining units it is being compared to. This is similar to the way of ranking athletes in a Decathlon competition in the Olympic Games, where each of the disciplines is punctuated separately and the athletes scores depend on their individual performance. Then, the scores are added up, and the athletes are ordered accordingly.

In some other disciplines a different approach is taken. For instance, in soccer or basketball a tournament is performed by confronting teams in pairs, and, following different rules, teams are ranked depending upon their performance in the different matches they participate. In the case of the soccer national leagues in Europe, all teams in the same division compete with each other twice, and the final ranking takes into account the result of all those pairwise matches. In tournaments, then, teams are not evaluated independently. Instead, the ranking depends both upon their performance against the competitors, and upon the competitors' performance. Because of that, to know whether one team is above another in the ranking it is not enough to know the results of the vis-à-vis matches between the two. Formally, the worth approach provides a *relative* evaluation of the U research units where the domain is the product space $R^{NI} \times \dots \times R^{NU}$, and the range is the non-negative orthant of R^U . As the evaluation depends upon the particular set of units under consideration, generally binary properties do not make sense –as in the case of properties **D** and **E**. On the other hand, when the property **F** of Independence –a necessary condition for any scoring rule (Marchant, 2009)– is

appropriately reformulated, it can be shown that the ranking induced by the worth vector does not satisfy it.

- The four indicators satisfy properties **A**, and **B**.
- The indicators *RCR*, *Topk*, and *ANGk* satisfy **D**, and hence **E** (for the role of these properties in citation analysis, see Perianes-Rodríguez & Ruiz-Castillo, 2015, 2016).

In what follows we comment on the pros and cons of the four types of indicators.

1. The next result characterizes the set of *Topk* indicators for any k .

Theorem 1. The class of subgroup consistent, scale invariant, and translation invariant high-impact indicators consist of monotonic transformations of the set of *Topk* indicators (For a discussion of this and the next characterization result, see Albarrán *et al.*, 2011)

Practical implications: if we choose a nontrivial high-impact indicator satisfying **B** and **C** that is not an increasing function of the *Topk* indicator, then the chosen indicator must violate subgroup consistency.

2. However, the double requirement of **B** and **C** can be meaningfully relaxed as follows. A scale invariant indicator and a translation invariant indicator are said to be *compatible* if, at any fixed CCL, they give the same ranking of distributions, although not necessarily the same values. A natural question to ask is if compatible pairs of indices exist that are subgroup consistent. The answer is the following.

Theorem 2. If a pair of subgroup consistent high-impact citation indicators is compatible, then the scale invariant index must be a monotonic transformation of a member of the FGT family defined in (3).

Therefore, if we choose to adopt a nontrivial high-impact indicator satisfying **B** that is not an increasing function of a member of the FGT family, then the chosen indicator must either violate continuity, subgroup consistency, or have no translation invariant high-impact index counterpart that ranks citation distributions in the same way for a given CCL. From this perspective, the choice of the set of *ANGk* indicators for any k is a reasonable one.

3. The main difference between *Topk* and *Worthk* indicators, on one hand, and *ANGk* indicators on the other, is that the former are not strictly monotonic on the set of high-impact articles. Consequently, we expect *ANGk* indicators to have the greatest discriminatory power.

4. As we have seen, in the dichotomous case *Worthk* indicators rank any set of research units exactly as *Topk* indicators. However, insofar as *Worthk* indicators take into account the direct and indirect relationships between all the units involved, we expect them to have a greater discriminatory power than *Topk* indicators.

5. The *RCR* indicator has many desirable properties. However, it is not able to focus on high-impact publications in the upper tail of citation distributions.

Data and empirical results

We begin with 4,472,332 distinct articles published in 1998-2003 with a five-year citation window for each year in that period. Each of these articles is assigned by Thomson Reuters to one of 22 broad fields. We consider 40 countries that have published at least 10,000 articles in all sciences in 1998-2003.

Among the counting methods that can be readily applied to co-authored articles, we follow a multiplicative strategy that extends as much as necessary the citation distributions of the research units in our dataset (For a recent discussion, see Waltman & Van Eck, 2015, and Perianes-Rodriguez & Ruiz-Castillo, 2015). In this way, we arrive at what we call the *geographical extended count* with 5,450,309 articles. Information concerning the distribution of articles by field in the original and the geographically extended count, as well as a detailed description the geographical areas are available on request.

For any indicator F_{ij} of unit i in field j , we construct the all-sciences indicator $F_i = \sum_j (N_{ij}/N_i)F_{ij}$, where $N_i = \sum_j N_{ij}$ is the total number of articles of unit i . This is exactly how *Topk* indicators are constructed in practice in the all-sciences case. When $F_{ij} = RCR_{ij}$, then $F_i = MNCS_i$. Whenever an indicator satisfies properties **A** and **B**, as in the case of the *Worthk* and *ANGk* indicators, this is an appropriate procedure of solving the all-sciences aggregation problem (Perianes-Rodriguez & Ruiz-Castillo, 2016). The results on the discriminatory power of the 10 indicators are in Table 2 (For reasons of space, results on country rankings are available on request).

Table 2. Discriminatory power (Range and CV) of ten citation impact indicators

	<i>MNCS</i>	<i>Top50%</i>	<i>Worth50</i>	<i>ANG50</i>	<i>Top10%</i>	<i>Worth10</i>	<i>ANG10</i>	<i>Top1%</i>	<i>Worth1</i>	<i>ANG1</i>
Range	1.093	0.743	1.116	1.120	1.285	1.362	1.545	1.690	1.696	1.897
CV	0.329	0.213	0.341	0.327	0.408	0.431	0.490	0.549	0.552	0.601

1. Among current indicators, the discriminatory power of the *MNCS* is greater than that of the *Top50%*. As expected, the discriminatory power increases as we focus on ever-smaller parts of the upper tail of citation distributions: $Top1\% > Top10\% > Top50\%$.

2. We confirm that *Worthk* indicators always have a greater discriminatory power than that of the corresponding *Topk%* indicators.

3. As expected, strict monotonicity implies that *ANGk* indicators always have the greatest discriminatory power: for every k , $ANGk > Worthk > Topk\%$.

Conclusions and further research

There is a broad agreement concerning which indicators we should use, namely, the *MNCS* among average-based indicators, and *Topk%* indicators for different k values among percentile rank indicators. This is a very reasonable choice: *Topk%* indicators allow us to focus on the upper tail of citation distributions and, in view of Theorem 1, they do have excellent properties.

However, this is not the end of the story. In the first place, we have introduced a new relative indicator with very good properties –the *Worth_k* indicator– that, in addition, exhibits a greater discriminatory power than *Topk%* indicators. In the second place, the *ANG_k* indicators, which are the only strictly monotonic indicators, exhibit the greatest discriminatory power.

The following questions are left for further research. First, there are several interesting non-dichotomous percentile rank indicators. For example, the canonical case in which the reference citation distribution is partitioned into 100 percentiles receiving scores from 1 to 100, or the indicator used by the National Science Foundation (2010). For each of them, a Worth indicator can be defined. Will Worth non-dichotomous indicators still exhibit a greater discriminatory power than the corresponding percentile rank indicators? Second, percentile rank indicators treat equally all publications within a given percentile or percentile class. Instead, Worth indicators can be applied to the distribution of distinct citation values in any scientific field. How would this possibility affect the discriminatory power of Worth indicators? Third, percentile rank indicators are supposed to be more robust to extreme observations than average-based indicators, including the *ANG_k* type discussed in this paper. However, Worth indicators should be as robust as the corresponding percentile rank indicators. It is important to verify these conjectures in practice.

REFERENCES

- Albarrán, P., Ortuño, I., & Ruiz-Castillo, J. (2011). The measurement of low- and high-impact in citation distributions: technical results. *Journal of Informetrics*, 5, 48–63.
- Bornmann, L., & Marx, W. (2013). How good is research really? *EMBO Reports*, 14, 226-230.
- Foster, J.E., J. Greer, & E. Thorbecke (1984). A Class of Decomposable Poverty Measures. *Econometrica*, 52, 761-766.
- Herrero, C., & Villar, A. (2013). On the comparison of group performance with categorical data. *PLoS ONE*, 8, e84784.
- Leydesdorff, L., Bornmann, L., Mutz, R., & Opthof, T. (2011). Turning the tables on citation analysis one more time: Principles for comparing sets of documents. *Journal of the American Society for Information Science and Technology*, 62, 1370-1381.
- Lunberg, J. (2007). Lifting the Crown –Citation z-score. *Journal of Informetrics*, 1, 145-154.
- Marchant, T. (2009). Score-based bibliometric rankings of authors. *Journal of the American Society for Information Science and Technology*, 60, 1132-1137.
- National Science Foundation (2010), *Science and Engineering Indicators 2010*, Appendix Tables, Arlington, VA, U.S.: National Science Foundation.
- Perianes-Rodriguez, A. & Ruiz-Castillo, J. (2015). Multiplicative versus fractional counting methods for co-authored publications. The case of the 500 universities in the Leiden Ranking. *Journal of Informetrics*, 9, 974-989.

- Perianes-Rodriguez, A. & Ruiz-Castillo, J. (2016). A comparison of two ways of evaluating research units working in different scientific fields. *Scientometrics*, 106, 539-561.
- Rousseau (2012). Basic properties of both percentile rank scores and the I3 indicator. *Journal of the American Society for Information Science*, 63, 416-420.
- Waltman, L., and Van Eck, N.J. (2009). A taxonomy of bibliometric performance indicators based on the property of consistency. Technical Report ERS-2009-014-LIS, Erasmus University Rotterdam, The Netherlands. *Presented at the 12th International Conference on Scientometrics and Informetrics*, Rio de Janeiro.
- Waltman, L., and Van Eck, N.J. (2015). Field normalized citation impact indicators and the choice of an appropriate counting method. *Journal of Informetrics*, 9, 872-894.
- Waltman, L, N. J. van Eck, T. N. van Leeuwen, M. S. Visser, & van Raan, A. F. J. (2011). Towards a New Crown indicator: Some theoretical considerations. *Journal of Informetrics*, 5, 37-47.



How does the scientific progress in developing countries affect bibliometric impact measures of developed countries? A counterfactual case study on China¹

Stephan Stahlschmidt* and Sybille Hinze**

* stahlschmidt@dzhw.eu

Department 2, German Centre for Higher Education Research and Science Studies (DZHW), Schützenstr. 6a, Berlin, 10117 (Germany)

** hinze@dzhw.eu

Department 2, German Centre for Higher Education Research and Science Studies (DZHW), Schützenstr. 6a, Berlin, 10117 (Germany)

ABSTRACT

Many developed countries have found their bibliometric impact measures to be improving over the last decade. Also the BRICS states, the economically largest group of developing countries, observe a similar pattern.

This uniform growth seems puzzling, as not every country can improve its relative performance to all other countries. A possible explanation for this uniform growth might be found in the dynamic environment and especially in the exponential growth of Chinese publications. We like to analyze how this unprecedented growth of contributions from a single country with its specific bibliometric characteristics affects the whole bibliometric measurement process.

We show that due to the lowly cited Chinese publications the overall corpus of scientific publications grows especially in the lower tail and argue that this unequal increase in publications benefits especially the bibliometric impact measures of developed countries. The actual magnitude of this effect will be derived by contrasting the actual bibliometric world with a counterfactual one without China.

MOTIVATION

Policy makers usually don't miss a good opportunity to show off the positive effects of their policies. Consequently the German government noticed with satisfaction the recent positive evaluation of the so-called "Exzellenzinitiative", a large government sponsored programme to strengthen research in Germany. It was especially noted that during the programme the bibliometric impact of German publications has improved, although the causal attribution of this increase to the "Exzellenzinitiative" could not be analysed (Hornbostel and Möller, 2015).

Germany is not the only country which has found its bibliometric impact measures to be improving over the last decade, but this increase denotes a common characteristic among most developed countries. Also the BRICS states, the economically largest group of developing countries, observe a similar pattern. Figure 1 and 2 depict the temporal progress of two of the most common bibliometric impact measures, the mean normalised citation score

¹ This work was supported by German Federal Ministry of Education and Research (Grant 01PQ13001)

(Figure 1) and the share of highly cited publications (Figure 2), for a group of developed and developing countries.

Only the United States (-0.2% p.a.) exhibit a negative growth rate in their MNCS in the observed time period 1995-2011, while no country has seen its share of highly cited publications falling in the same time period. Consequently countries have improved the performance in these relative measures of bibliometric impact with China observing the largest growth rate in MNCS (3.9% p.a.) and the second largest in share of highly cited publications (24.8% p.a.).

This uniform growth among developed and the largest developing countries seems puzzling: The listed bibliometric impact measures are relative measures and, other things being equal, not every country can improve its relative performance to all other countries, but consequently some, fairly large, countries must be affected by a declining bibliometric impact.

Figure 1: Mean normalised citation score by country (wholecount).

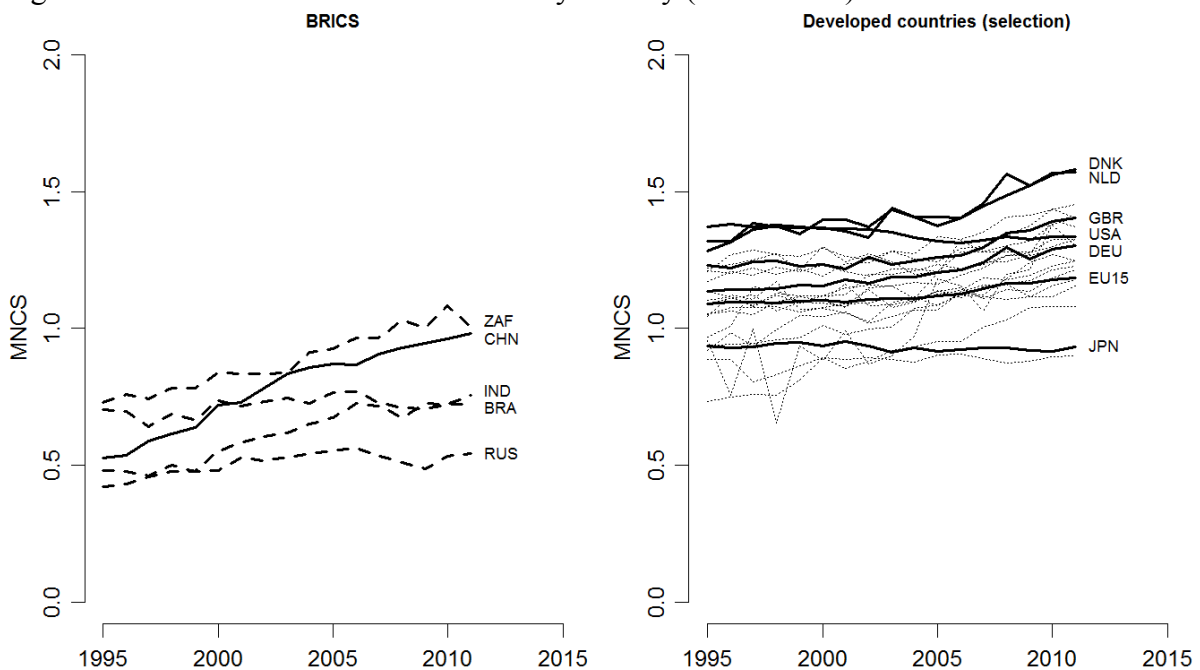
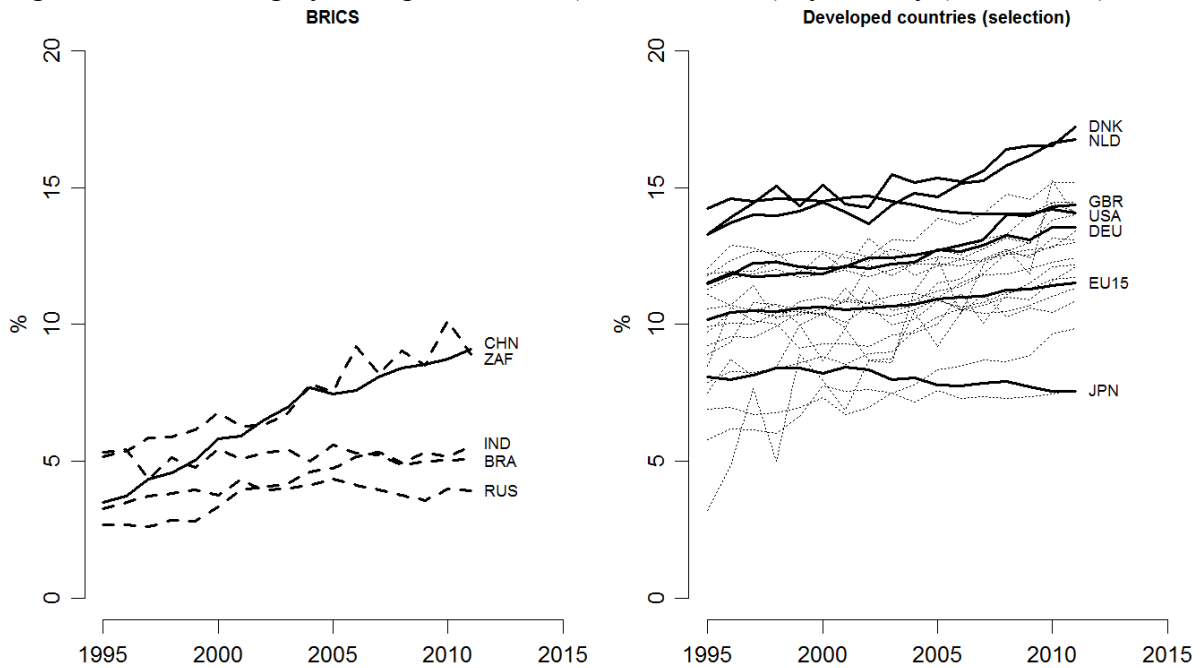
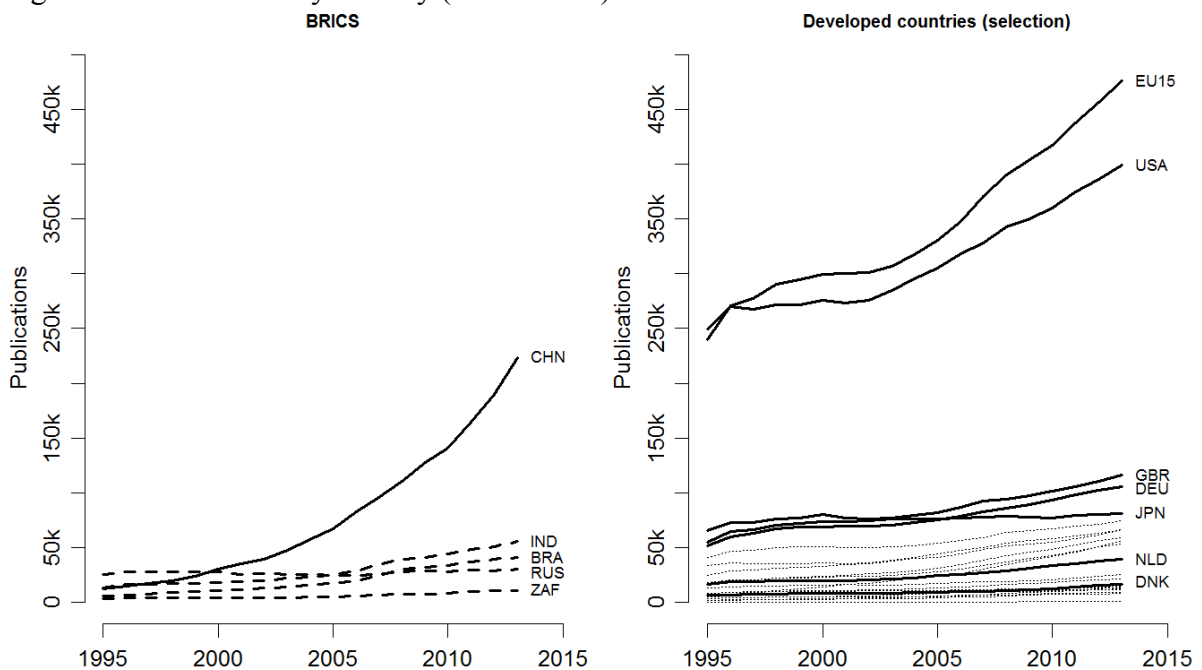


Figure 2: Share of highly cited publications (>90% Quantile) by country (wholecount).



A possible explanation for this uniform growth might be found in the dynamic environment in which these relative measures are embedded. As depicted in Figure 3 the number of articles, reviews and letters in journals indexed by Thomson's Reuter's Web of Science increases modestly by about 4% p.a., while the contribution of China to this set has been growing exponentially. Its annualised growth rates tops 17% p.a. surpassing all other BRICS countries (Brazil: 11.8% p.a., India: 8.2% p.a., South Africa: 6.4% p.a. and Russia: 0.8% p.a.) and developed countries (e.g. USA: 2.6% p.a and EU15: 3.9% p.a.).

Figure 3: Publications by country (wholecount)².



² Countries without label (grey lines): Australia, Canada, EU15 member states, Israel, New Zealand, Republic of Korea

Sine 2006 China counts as the second biggest single “producer” of publications trailing only the United States and the question arises, how this unprecedented growth of contributions from a single country with its specific bibliometric characteristics affects the whole bibliometric measurement process.

While former work on the topic is often concerned with describing the impact of the Chinese appearance at the forefront of scientific publications, e.g. country shares of publications/citations (Côté et al 2016, Bornmann and Leydesdorff, 2013; Leydesdorff and Wagner, 2009) or relating these bibliometric measures with socio-economic data (May, 1997), we are more interested in the arising measurement issues implied by the growth of Chinese publications. Consequently we will contrast the just described state of the world of scientific publications with a counterfactual bibliometric world without China. We'd like to analyse how developed countries would have performed under stable conditions and investigate if and how the increase in Chinese publications is related to the uniform growth of bibliometric impact measures among developed countries described above.

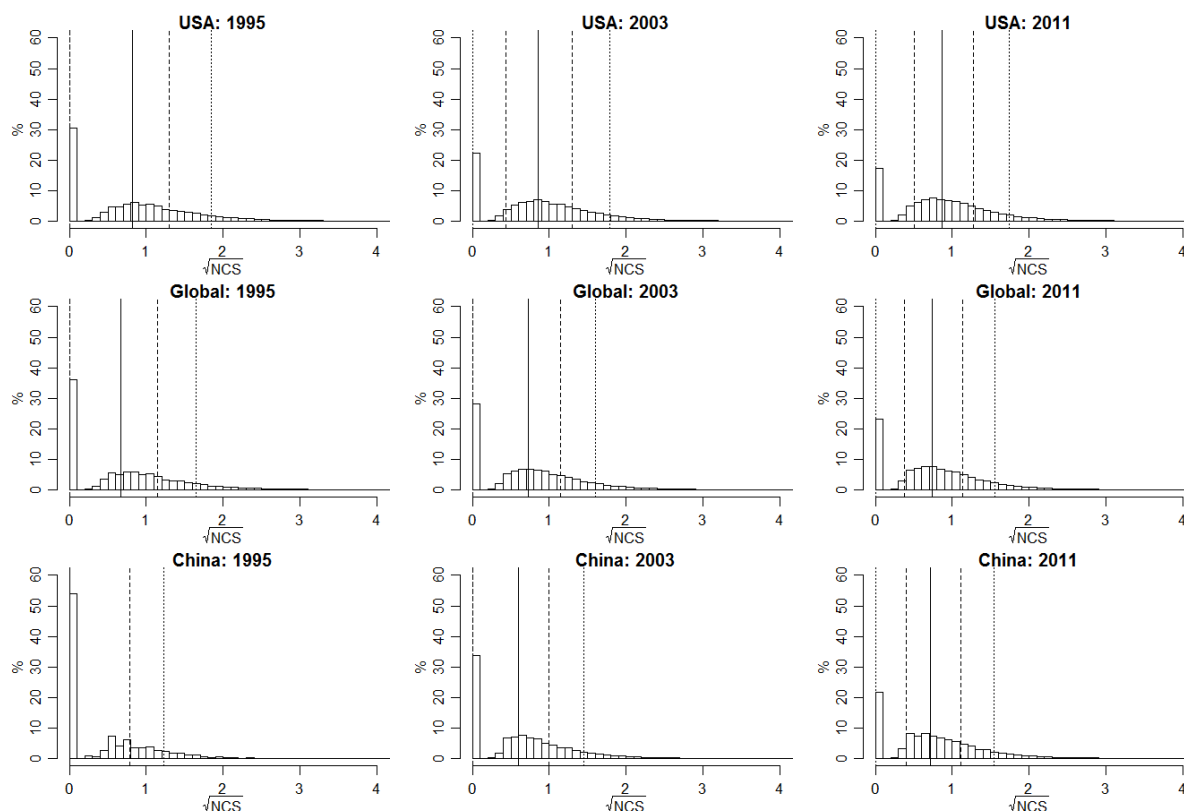
This research-in-progress comprises several implications for policy setting agents. We like to answer, to what extent the observed increase in bibliometric impact measures in the developed world might ultimately be assigned to country specific science programmes and policies and to what extent the improvement denotes an artificial measurement artefact caused by a dynamic environment and respectively the way the indicators are constructed which are used to reflect a country's scientific performance.

BIBLIOMETRIC IMPACT MEASURES AND THE CITATION DISTRIBUTION

The observed exponential growth of Chinese publications might affect bibliometric impact measures only if its citation distribution differs from the worldwide citation distribution. If not, the Chinese publications only add further publications to the corpus of scientific publications allowing for more precision in the measurement process.

Figure 4 depicts in the last row the citation distribution of Chinese publications, i.e. the year and subject field specific normalised citation score of every article, review or letter appearing in a Web of Science indexed journal based on triannual citation window. According to this illustration China has come a long way starting with median normalised citation score of 0 in 1995 to 0.51 in 2011. At the same time the 10% best Chinese publications achieved a normalised citation score of 1.5 in 1995, respectively 2.4 in 2011.

Figure 4: Histogram, median (solid line), 25% and 75 % quartiles (dashed lines) and 10% and 90% quantiles (dotted lines) of the normalised citation score for American, Chinese and worldwide (sample) publications in the years 1995, 2003 and 2011. Squared scale is displayed for graphical convenience and data is truncated in the upper tail.



Comparing the distribution of Chinese publications with the American or worldwide publications, it might be observed, that the Chinese publications exhibit a specific citation pattern. This pattern differs largely from the developed countries and worldwide equivalents and possesses a pronounced lower tail with relatively numerous publications obtaining zero or few citations. A two-sample Kolmogorow-Smirnow-Test (unreported) rejects the Null-hypothesis that the Chinese and either American or worldwide publications are drawn from the same super-population. Consequently it seems reasonable to assume that due to the inclusion of Chinese publication, the corpus of scientific publications grows especially in the lower tail.

At the same time the currently used bibliometric impact measures are relative indicators and factor the whole distribution, including the growing lower tail, in. Naturally this extension of the publication base affects also the bibliometric impact measures MNCS and share of highly cited publications. In detail, the appearance of many low cited publications lowers the field specific expected citation count and the MNCS ratio of actual citations to expected citations will increase due to the smaller expected count. On the other hand, the specific Chinese citation pattern with many relatively low cited publications will result in an uneven expansion of the publication corpus and disproportionally reduce the 90% quantile.

Hence the arrival of the Chinese publications lowers the relative thresholds (expected count, 90% quantile) of the corresponding bibliometric impact measures. Publications in the upper tail will consequently thrive upon this shift of the relative thresholds, as they will either approach from below, surpass or even outdistance these newly set thresholds. Although the citation count of highly cited publications stays constant (or at best increases due to citations from the new Chinese arrivals), they improve their relative standing to the lowered thresholds. Especially developed countries, which in general find a larger share of their publications in the upper tail of the worldwide citation distribution, should benefit from this mechanism.

Still the Chinese publications might not reverse the overall evolution of the citation corpus. For example even with the Chinese publications included the median (over subject fields) expected citation count has still doubled from 1.49 to 2.93 between 1995 and 2011. Hence the arrival of the Chinese low cited publications might rather (temporary) slow down the growth of the thresholds, but does not seem pronounced enough to upturn persisting trends.

OUTLOOK: A COUNTERFACTUAL WORLD WITHOUT CHINA

In the next step, we would like to construct a counterfactual bibliometric world without China. Therefore an alternative publication corpus will be derived, which excludes all institutions from China, and the normalised citation scores and share of highly cited publications will be computed for the remaining countries. Obviously this counterfactual world will denote only an approximation, as the Chinese science system is still partially included via Chinese researcher working in non-Chinese institutes.

Leaning on the treatment effect literature in economics (Imbens and Wooldridge, 2009) we will match each country to its counterfactual twin and contrast the actual and counterfactual bibliometric impact measures to gain insights into the consequences of the Chinese publication growth on bibliometric impact measure of developed countries.

Issues we will consider in our analyse include the inclusion policy of new journals in the WoS index, non-English language as a citation barrier, the modelling of the country specific citation pattern with zero-inflated distributions and the implications of fractional counting.

REFERENCES

- Bornmann L. and Leydesdorff, L. (2013) Macro-Indicators of Citation Impacts of Six Prolific Countries: InCites Data and the Statistical Significance of Trends. *PLOS ONE*. DOI: 10.1371/journal.pone.0056768
- Côté, G., Roberge, G. and Archambault, E. (2016) Bibliometrics and Patent Indicators for the Science and Engineering Indicators 2016. Retrieved March 30, 2016 from http://www.science-metrix.com/files/science-metrix/publications/science-metrix_comparison_of_2016_bibliometric_indicators_to_2014_indicators.pdf
- Hornbostel, S. and Möller, T. (2015). *Die Exzellenzinitiative und das deutsche Wissenschaftssystem. Eine bibliometrische Wirkungsanalyse*. Retrieved March 29, 2016 from http://www.bbaw.de/publikationen/wissenschaftspolitik_im_dialog/BBAW_WiD-12_PDF-A1b.pdf
- Imbens, G. and Wooldridge, J (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, 47, 5-86.
- Leydesdorff, L. and Wagner, C. (2009). Is the United States losing ground in science? A global perspective on the world science system. *Scientometrics*, 78, 23-26.
- May, Robert M. (1997). The scientific wealth of nations. *Science*, 275, 793-796.



A comparison of the Web of Science with publication-level classification systems of science

Antonio Perianes-Rodríguez* and Javier Ruiz-Castillo**

* antonio.perianes@uc3m.es

Departamento de Biblioteconomía y Documentación, Universidad Carlos III, SCImago Research Group

** jrc@eco.uc3m.es

Departamento de Economía, Universidad Carlos III

ABSTRACT

In this paper we propose a new criterion for choosing between a pair of classification systems of science that assign publications (or journals) to a set of scientific fields. Consider the standard normalization procedure in which field mean citations are used as normalization factors. We recommend system A over system B whenever the standard normalization procedure based on A performs better than the when it is based on B. Since the evaluation can be made in terms of either system, the performance assessment requires a double test. In addition, since the assessment of two normalization procedures would be generally biased in favor of the one based on the classification system used for evaluation purposes, ideally a pair of classification systems must be compared using a third, independent classification system for evaluation purposes. We illustrate this strategy by comparing a Web of Science journal-level classification system, consisting of 236 journal subject categories, with two publication-level algorithmically constructed classification systems consisting of 1,363 (G6) and 5,119 (G8) clusters. There are two main findings. (1) The G8 system is found to dominate the G6 system. Therefore, when we have a choice between two classification systems at different granularity levels, we should use the system at the higher level because it typically exhibits a better standard normalization performance. (2) The G8 system and the Web of Science (WoS) journal-level system are found to be non-comparable. Nevertheless, the G8-normalization procedure performs better using the WoS system for evaluation purposes than the WoS-normalization procedure using the G8 system for evaluation purposes. Furthermore, when we use the G6 system for evaluation purposes, the G8-normalization procedure performs better than the WoS-normalization procedure. We conclude that algorithmically constructed classification systems constitute a credible alternative to the WoS system and, by extension, to other journal-based classification systems.

ACKNOWLEDGEMENTS. This research project builds on earlier work started by Antonio Perianes-Rodríguez during a research visit to the Centre for Science and Technology Studies (CWTS) of Leiden University as awardee of José Castillejo grant, CAS15/00178, funded by the Spanish MEC. Ruiz-Castillo is a visiting researcher at CWTS and gratefully acknowledges CWTS for the use of its data. Ruiz-Castillo acknowledges financial support from the Spanish MEC through grant ECO2014-55953-P, as well as grant MDM 2014-0431 to his Departamento de Economía.

INTRODUCTION

For many theoretical and practical purposes in the evaluation of research activities in current society, we need a *classification system* of science, that is, an assignment of individual publications (or journals) to a set of clusters or sub-fields. As is well known, the choice of a classification system remains an open question in Scientometrics (see *inter alia* Small, 1999, Boyack *et al.*, 2005, Leydesdorff, 2004, 2006, and Leydesdorff and Rafols, 2009, as well as the references they contain). Together with the classification systems included in Thomson Reuters' Web of Science (WoS hereafter) and Elsevier's Scopus databases, there are a number of interesting proposals suggested by individual researchers (see *inter alia* Börner *et al.* (2012), as well as the references in Waltman & Van Eck, 2012).

In this paper, we contribute to the search for an appropriate classification system begun in Ruiz-Castillo & Waltman (2015). The main idea is the following. Given a classification system, it is well known that differences in production and citation practices preclude the direct comparison of the raw citations received by any pair of publications belonging to different clusters. In this situation, one way to evaluate the performance of research units working in different clusters begins with the normalization of the original citation counts. Given its simplicity and good results (Radicchi *et al.*, 2008, Li *et al.*, 2013), consider the standard target (or cited-side) normalization procedure in which normalized citation scores in every cluster are equal to the original raw citations divided by the cluster mean citation. If one could establish that the standard normalization procedure based in system A performs better—in a sense made precise below—than the standard normalization procedure based in system B, then we would recommend the use of system A over system B.

We illustrate this strategy by comparing a Web of Science (WoS hereafter) journal-level classification system, consisting of 236 journal subject categories (or simply categories hereafter), with two alternatives arising from the publication-level algorithmic methodology introduced in Waltman & Van Eck (2012) that classifies individual publications into clusters solely based on direct citations between them.

In practice, the choice of the WoS classification system is often made because, together with the Scopus system, it is readily available. However, a number of studies question the appropriateness of this system for normalization purposes (Ruiz-Castillo & Waltman, 2015). Among the publication-level alternatives, Klavans and Boyack (2015) conclude that classification systems based on direct citation using the Waltman & Van Eck (2012) methodology are more accurate than classification systems based on bibliographic coupling or co-citation. Ruiz-Castillo & Waltman (2015) apply the publication-level algorithmic methodology introduced by Waltman & Van Eck (2012) to a WoS dataset consisting of 9.4 million publications from the 2003-2012 period. They construct a sequence of twelve independent classification systems, in each of which the same set of publications is assigned to an increasing number of clusters.

Which granularity level is used in practice in the calculation of normalized citation impact indicators is a very important issue (Zitt *et al.*, 2005). Taking into account the joint variation of cluster size and cluster mean citation in the different classification systems, the degree of skewness and the similarity of this characteristic across cluster citation distributions, and the degree of homogeneity within cluster citation distributions, Ruiz-Castillo & Waltman (2015) recommend using granularity levels 7 and 8. To emphasize the potential differences between granularity levels, in this paper we select levels 6 and 8 (the

G6 and G8 classification systems hereafter) consisting of 1,363 and 5,119 clusters, respectively. Therefore, we focus on the following two comparisons: the G6 versus the G8 system, and the winner in this contest versus the WoS system.

We focus on the 3.6 million articles published in the 2005-2008 period, and the citations they receive during a five-year citation window for each year in that period. We adopt a multiplicative strategy to solve the problem that approximately 45% of the articles in the WoS system are assigned to several categories. The total number of articles in what we call the *extended count* for the 236 WoS categories is 5.9 million, or 64.5% larger than the original dataset. On the other hand, since the methods for the evaluation of normalization procedures in Li & Ruiz-Castillo (2013) require the partition of cluster (and category) citation distributions into, say, 100 quantiles, we eliminate clusters (and categories) with less than 250 articles. In the G8 system we are left with 3.4 million articles that are classified into 3,332 clusters, while in the G6 system we are left with 3.6 million articles that are classified into 900 clusters. For comparison purposes, we also eliminate the five WoS categories with less than 250 articles.

The comparison between the G6 and G8 classification systems

The assessment of the performance of the standard field-normalization procedure when it is based on different classification systems

How can we compare the performance of the standard field-normalization procedure when it is based in two different classification systems? A possible answer is to use the graphical and numerical methods introduced in Li & Ruiz-Castillo (2013) for that purpose. To save space, in this paper we only report our findings under the graphical approach. A discussion and results using the numerical approach can be found in the Working Paper version of the paper, Perianes-Rodriguez & Ruiz-Castillo (2016), or PRRC hereafter.

The graphical approach

In this sub-section, we evaluate the performance of the G6 and G8 classification systems using a method that applies the additive decomposability property of a certain member of the Generalized Entropy family of citation inequality indices –denoted by I – to the double partition of the data into clusters and quantiles (see Crespo *et al.*, 2013, for details). Consider, for example, the case of the G8 system. Partition each cluster citation distribution into its 100 percentiles, denoted by $\pi=1, \dots, 100$. The citation inequality between clusters at each percentile, denoted by $I(\pi)$, is entirely attributable to the differences in citation practices between the 3,332 clusters holding constant the degree of excellence in all clusters at percentile π . Therefore, the performance of any normalization procedure can be assessed in terms of its ability to reduce the terms $I(\pi)$ at every π . Consequently, to assess the impact of the G8-normalization procedure using the G8 system for evaluation purposes, we simply observe how expressions $I(\pi)$ vary when we compute them for the normalized citation distributions. The two alternatives, before and after normalization, correspond to the blue and the green lines in Figure 1 (Since the terms $I(\pi)$ are very high for percentiles in the lower tail of citation distributions, for clarity all Figures in this paper only include percentiles π in the interval [46, 100]). Note that the impact of the G8-normalization procedure is very important: the green line is considerably below the blue line at all percentiles.

For the comparison of the G8- and G6-normalization procedures, we extend the methods introduced in Li & Ruiz-Castillo (2013) to take into account that these classification systems have a different number of articles (details are available in PRRC). The orange line in Figure 1 represents the expressions $I(\pi)$ for the G6-normalization procedure when the G8 system is used for evaluation purposes. The fact that the orange line is above the green line indicates that the effect of the G6-normalization procedure in reducing $I(\pi)$ at every π is not as strong as the effect of the G8-based normalization procedure. We say that the latter uniformly dominates the former using the G8 classification system for evaluation purposes.

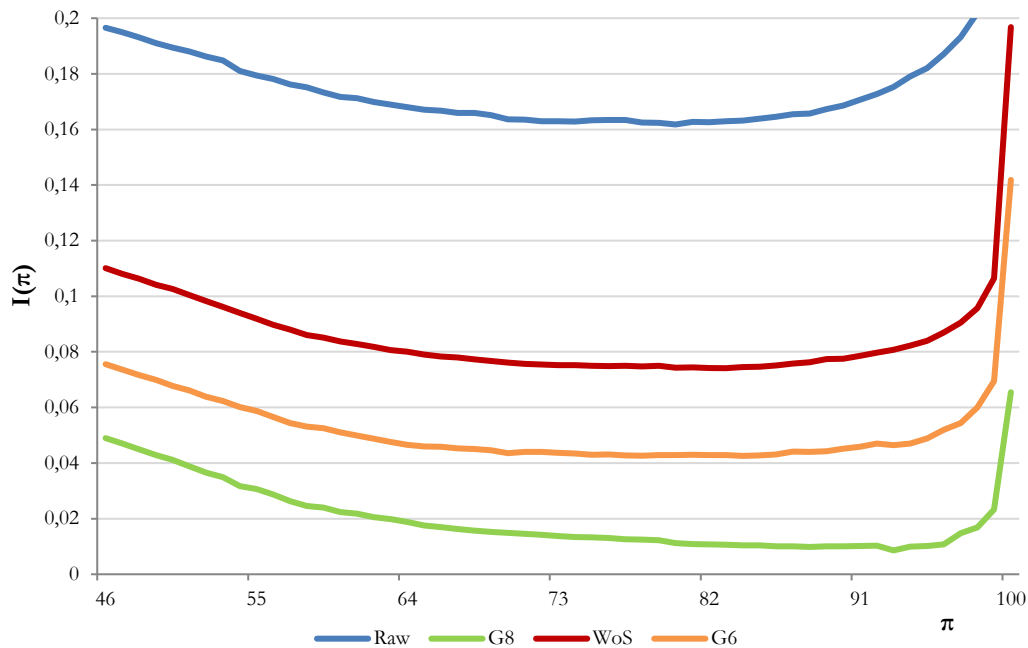


Figure 1. $I(\pi)$ terms for percentiles in the interval $[46, 100]$ before and after normalization using the G8 classification system for evaluation purposes

It has been argued that the assessment of two normalization procedures would be generally biased in favor of the normalization procedure based on the classification system used for evaluation purposes (Sirtes, 2012, and Waltman & Van Eck, 2013). According to Li & Ruiz-Castillo (2013), a solution consists of adding a second test to the above procedure where the G6 system is now used for evaluation purposes. The two alternatives, before and after normalization, correspond to the blue and the orange lines in Figure 2. Again, the impact of the G6-normalization procedure is very important: the orange line is always clearly below the blue line for all π .

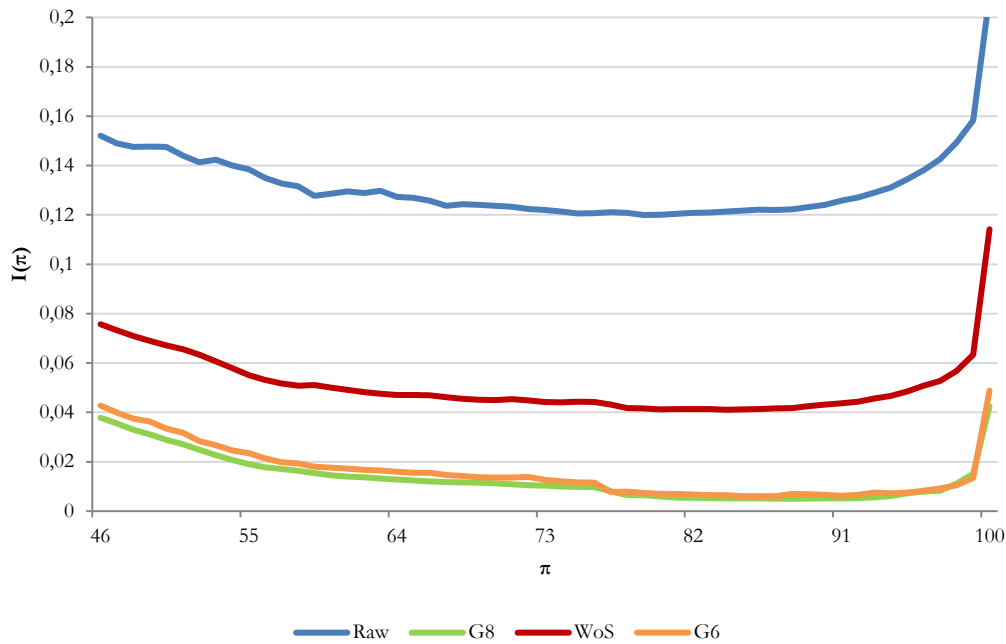


Figure 2. $I(\pi)$ terms for percentiles in the interval $[46, 100]$ before and after normalization using the G6 classification system for evaluation purposes

The green line in Figure 2 represents the expressions $I(\pi)$ for the G8-based normalization procedure when the G6 system is used for evaluation purposes. Since the green and the orange line intersect at some percentiles, the G8- and the G6-normalization procedures are non-comparable when using the G6 system for evaluation purposes. However, since the former uniformly dominates the latter using G8 as the evaluation classification system, we conclude that the G8-normalization procedure weakly dominates the G6-normalization procedure according to the double test under the graphical approach.

Robustness analysis

Ideally, for comparing two normalization procedures based in two different classification systems we should use a third, independent system, for evaluation purposes (Sirtes, 2012, and Waltman & Van Eck, 2013). We find that, at least in the upper tails of clusters' and categories' citation distributions, the systems G6 and G8 are quite different from the WoS system (for details, see PRRC). Therefore, we suggest comparing the G6- and G8-normalization procedures using the WoS classification system for evaluation purposes. The expressions $I(\pi)$ corresponding to the organization of the raw data are represented by the blue line in Figure 3, while the orange and the green lines represent the effect of the G6- and the G8-normalization procedures. Clearly, we conclude that the G8 system strictly dominates the G6 system under the graphical approach using the WoS classification system for evaluation purposes.

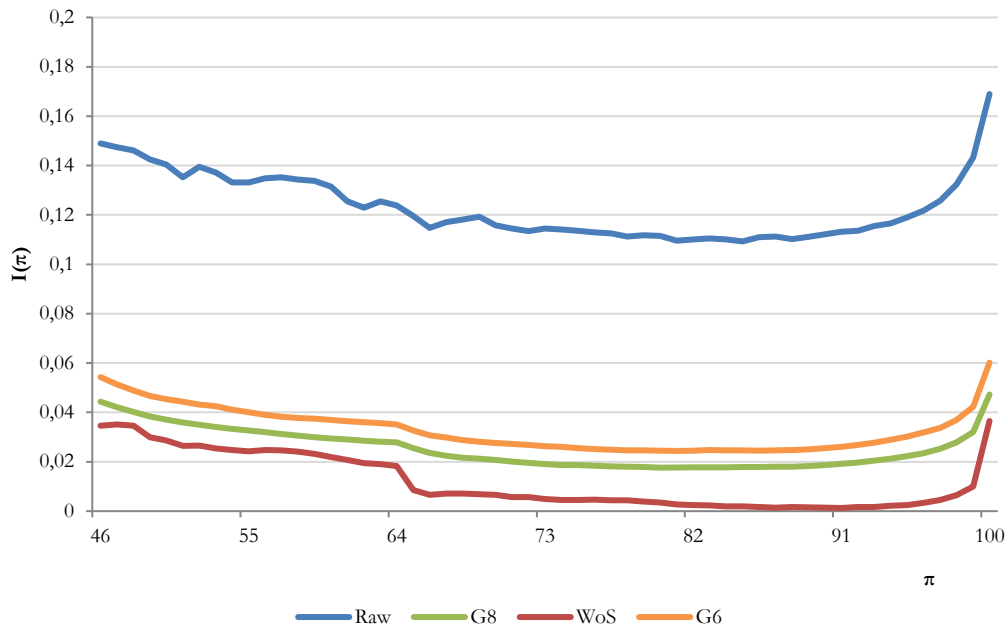


Figure 3. $I(\pi)$ terms for percentiles in the interval $[46, 100]$ before and after normalization using the WoS classification system for evaluation purposes

The comparison between the G8 and WoS classification systems

We begin by assessing the performance of the WoS-normalization procedure when the G8 system is used for evaluation purposes. The red line in Figure 1 represents the expressions $I(\pi)$ for the WoS-based normalization procedure when the G8 system is used for evaluation purposes. The fact that the red line is above the green line indicates that the effect of the WoS-based normalization procedure in reducing $I(\pi)$ at every π is not as strong as the effect of the G8-based normalization procedure. We say that the latter uniformly dominates the former using G8 as the evaluation classification system.

Next, we must assess the performance of the G8- and the WoS-normalization procedures using the WoS system for evaluation purposes. The two alternatives, before and after normalization, correspond to the blue and the red lines in Figure 3. It is observed that the impact of this normalization is very important: the red line is always clearly below the blue line for all π . In turn, recall that the green line in Figure 3 represents the expressions $I(\pi)$ for the G8-normalization procedure when the WoS system is used for evaluation purposes. Since the red line is below the green line for all π , the effect of the WoS-normalization procedure in reducing $I(\pi)$ at every π is stronger than the effect of the G8-normalization procedure. We say that the former uniformly dominates the later using the WoS classification system for evaluation purposes.

We conclude that, in terms of this double test, the two normalization procedures are non-comparable under the graphical approach. Nevertheless, insofar as in Figure 3 the distance between the green and the red lines is smaller than in Figure 1, we may say that the G8-normalization procedure performs better using the WoS system for evaluation purposes than the WoS-normalization procedure using the G8 system for evaluation purposes.

Robustness analysis

Although differences between the WoS and the G6 systems are greater than differences between the G8 and the G6 systems, the latter are by no means negligible (see PRRC). Consequently, we believe that it is useful to compare the G8- and WoS-normalization procedures using the G6 classification system for evaluation purposes. The green line in Figure 2 represents the effect of the G8-normalization procedure using the G6 system for evaluation purposes. In turn, the red line in Figure 2 represents the effect of the WoS-normalization procedure using the G6 system for evaluation purposes. The fact that the red line is always above the green line in Figure 2 indicates that the G8-normalization procedure strongly dominates the WoS-normalization procedure when using G6 as the evaluation classification system.

DISCUSSION AND CONCLUSIONS

Our main results are the following two.

1. The possibility that using a classification system for evaluation purposes bias the analysis in favor of the normalization procedure based in this system, makes very difficult to conclude that one system-based normalization procedure overcomes another according to the double tests in the graphical and the numerical approaches. This is why the following finding is remarkable: system G8 dominates system G6 both in the graphical approach. In addition, when the WoS system is used for evaluation purposes, the G8 system graphically dominates system G6.

These results have important practical consequences. Firstly, when we have a choice between two classification systems at different granularity levels, we should use the system at the higher level because it typically exhibits a better standard normalization performance when cluster mean citations are used as normalization factors. Secondly, the G6-normalization procedure has been found to perform well not only under the G6 system itself, but also when its performance is assessed using the G8 or the WoS systems for evaluation purposes. Therefore, if there is only available a single classification system at an appropriately high granularity level, we should use it in the knowledge that the reduction of the effect on overall citation inequality attributable to differences in production and citation practices –even at higher granularity levels– is non-negligible.

2. The choice of an adequate classification system constitutes a problem for which there is no perfect solution: all options involve a certain degree of arbitrariness in the way clusters are selected. Nevertheless, using a set of new gold standards –consisting of articles with at least 100 references–, Klavans & Boyack (2015) compare publication-level algorithmically constructed classification systems based in direct citations à la Waltman & Van Eck (2012) with six journal-level systems that do not include the WoS. They conclude that the former are more accurate than the latter in the sense that they are better at concentrating references. Furthermore, it can be argued that publication-level systems are better able to handle publications in multidisciplinary journals and in other journals with a broad scope, and can be expected to offer an up-to-date representation of the structure of scientific fields (Ruiz-Castillo & Waltman, 2015). On the other hand, it should be recognized that algorithmically constructed classification systems at sufficiently high granularity levels pose a troublesome labeling problem that, in certain contexts, may limit its applicability.

In this context, this paper has compared the G8- and WoS-based standard normalization procedures. The main result is that, according to the graphical approach the two procedures are non-comparable. Nevertheless, the G8-normalization procedure performs better using the WoS system for evaluation purposes than the WoS-normalization procedure using the G8 system for evaluation purposes. Furthermore, when we use the G6 system for evaluation purposes, the G8-normalization procedure performs better than the WoS-normalization procedure in the graphical and numerical sense.

We conclude that algorithmically constructed classification systems constitute a credible alternative to the WoS system and, by extension, to other journal-based classification systems. Consequently, we celebrate the decision by the Centre for Science and Technology Studies of adopting an algorithmically constructed classification system of this type consisting of 3,822 clusters in the 2015 edition of the Leiden Ranking.

REFERENCES

- Börner K, Klavans R, Patek M, Zoss AM, Biberstine JR, Light RP, et al. (2012) Design and Update of a Classification System: The UCSD Map of Science. *PLoS ONE* 7(7): e39464.
- Boyack, K., R. Klavans, and K. Börner (2005). Mapping the backbone of Science. *Scientometrics*, 64, 351-374.
- Crespo J A, Li Y, Ruiz-Castillo J. (2013). The measurement of the effect on citation inequality of differences in citation practices across scientific fields. *PLoS ONE*, 8, e58727.
- Klavans R and Boyack KW. (2015). *Which type of citation analysis generates the most accurate taxonomy of scientific and technical knowledge?* [arXiv:1511.05078v1](https://arxiv.org/abs/1511.05078v1) [cs.DL].
- Leydesdorff, L. (2004). Top-down Decomposition of the Journal Citation Report of the Social Science Citation Index: Graph- and Factor Analytical Approaches. *Scientometrics*, 60, 159-180.
- Leydesdorff, L. (2006). Can Scientific Journals Be Classified in Terms of Aggregated Journal-Journal Citation Relations Using the Journal Citation Reports?. *Journal of the American Society for Information Science and Technology*, 57, 601-613.
- Leydesdorff, L. and I. Rafols (2009), “A Global Map of Science Based on the ISI Categories”, *Journal of the American Society for Information Science and Technology*, 60: 348-362.
- Li Y, Castellano C, Radicchi F, Ruiz-Castillo J. (2013). Quantitative evaluation of alternative field normalization procedures. *Journal of Informetrics*, 7, 746–755.
- Li Y and Ruiz-Castillo J. (2013). The comparison of normalization procedures based on different classification systems. *Journal of Informetrics*, 7, 945-958.

- Perianes-Rodríguez A and Ruiz-Castillo J. (2016). *A comparison of the Web of Science with publication-level classification systems of science*. Working Paper, 16-02. Madrid: Universidad Carlos III (<http://hdl.handle.net/10016/22137>).
- Radicchi F, Fortunato S, Castellano C. (2008). Universality of Citation Distributions: Toward An Objective Measure of Scientific Impact. *PNAS*, *105*, 17268-17272.
- Ruiz-Castillo J and Waltman L. (2015). Field-normalized citation impact indicators using algorithmically constructed classification systems of science. *Journal of Informetrics*, *9*, 102-117.
- Sirtes D. (2012), Finding the Easter Eggs Hidden by Oneself: Why Radicchi and Castellano's (2012) fairness test for citation indicators is not fair, *Journal of Informetrics*, *6*: 448– 450.
- Small, H. (1999). Visualizing Science by Citation Mapping. *Journal of the American Society for Information Science*, *50*, 799-813.
- Waltman L and Van Eck NJ. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology*, *63*, 2378–2392.



Ranking Journals Using Social Choice Theory Methods: A Novel Approach in Bibliometrics¹

Fuad T. Aleskerov², Vladimir Pisyakov³, Andrey N. Subochev⁴

ABSTRACT

We use data on economic, management and political science journals to produce quantitative estimates of (in)consistency of evaluations based on seven popular bibliometric indicators (impact factor, 5-year impact factor, immediacy index, article influence score, h-index, SNIP and SJR). We propose a new approach to aggregating journal rankings: since rank aggregation is a multicriteria decision problem, ordinal ranking methods from social choice theory may solve it. We apply either a direct ranking method based on majority rule (the Copeland rule, the Markovian method) or a sorting procedure based on a tournament solution, such as the uncovered set and the minimal externally stable set. We demonstrate that aggregate rankings reduce the number of contradictions and represent the set of single-indicator-based rankings better than any of the seven rankings themselves.

INTRODUCTION

After almost a century since Gross and Gross published their pioneering work (1927), ranking journals remains a problem. Introduction of the impact factor by Garfield and Sher (1963) ushered in the era of indicators. The emergence of the Scopus database and invention of the h-index (Hirsch 2005) reignited interest in developing various bibliometric measures. However, their growing multiplicity generates two questions:

- (a) How do rankings based on different measures correlate with each other?
- (b) How can we construct a “harmony” of rankings?

To answer the first question, we apply rank correlation analysis to rankings based on seven popular indicators. We find that all rankings positively correlate with each other, but there is a percentage of contradictions. We see no sufficient reason to presume that any indicator is somehow inferior to others. Therefore instead of trying to choose “the best” indicator, we suggest pooling the information contained in all rankings, even though this information is

¹ The study was financially supported through the Basic Research Program at the National Research University Higher School of Economics (HSE) and by the Russian Academic Excellence Project '5-100'.

² DeCAN Lab and Department of Mathematics, Faculty of Economic Sciences, National Research University Higher School of Economics, Moscow, Russia; Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia; alesk@hse.ru

³ Library, National Research University Higher School of Economics, Moscow, pisyakov@hse.ru

⁴ DeCAN Lab and Department of Mathematics, Faculty of Economic Sciences, National Research University Higher School of Economics, Moscow, asubochev@hse.ru

contradictory. For this purpose, we propose to use ordinal aggregation methods originated in social choice theory. To the best of our knowledge, these methods have never been used to rank journals. Rank correlation analysis confirm that aggregate rankings reduce the number of contradictions and represent the set of single-indicator-based rankings better than any of the seven rankings themselves.

DATA

We consider three sets of journals representing three academic disciplines: economics, management and political science. Rankings are computed for each set separately. Sets of journals were taken from Journal Citation Reports database from Thomson Reuters, along with their IF, 5-year IF, immediacy index and AI indicators (all for JCR-2011 edition). SNIP and SJR metrics for 2011 were taken from Journal Metrics website powered by Scopus database; h-index for each journal was calculated manually by searching Web of Science database. To make h-index more definite, the exact publication and citation windows have been applied. Only papers appeared from 2007 to 2011 have been considered, and citations to them made during the same period, 2007–2011.

The selection of indicators contains all kinds of metrics. There are un-weighted as well as weighted (AI, SJR) measures. Indicators use different publication windows, from one (immediacy index) to five (5-year IF, AI) years. Moreover, they are taken from different databases. A choice of a database may significantly change the values of indicators even when they are based on the same methodology (Pislyakov 2009). Data sources and properties of metrics are summarized in Table 1.

Table 1. Indicators: sources and properties.

	<i>Database</i>	<i>Year</i>	<i>Publication window, years</i>	<i>Weighted</i>
2-year IF	WoS/JCR	2011	2	No
5-year IF	WoS/JCR	2011	5	No
immediacy index	WoS/JCR	2011	1	No
article influence	WoS/JCR	2011	5	Yes
h-index	WoS	2007–2011 (papers and citations)	5	No
SNIP	Scopus	2011	3	No
SJR	Scopus	2011	3	Yes

After exclusion of publications with missing values, the sets contain 212 economic journals, 93 management science journals and 99 political science journals.

METHODS

We consider ranking of journals as a multicriteria decision problem. A classical solution is to apply some aggregation function, for instance a weighted sum, to alternative's criterial values and then rank alternatives by respective values of the function. However, this method has a fundamental deficiency related to its *cardinal* nature. To obtain meaningful results, one has to be sure that all aggregated indicators admit meaningful inter-indicator comparisons. In economics, this problem is known as the problem of interpersonal comparability of utilities. Bergson, Samuelson and Little built the so-called “new” welfare economics upon a postulate of incomparability of individual utilities. Arrow, the father of social choice, adopted this postulate and developed an *ordinal* approach to the aggregation problem (Arrow 1951). We propose to apply ordinal ranking methods from social choice since they are immune to incomparability problem and it is possible to frame any multicriteria decision problem as a social choice problem (Arrow & Raynaud 1986).

Basic notions

One of the main objectives of social choice theory is to determine what alternatives will be or should be chosen given a set of feasible alternatives and preferences of decision-makers (voters, experts). It is possible to transfer social choice methods to a multi-criteria setting if one treats a ranking based on a certain criterion as a representation of preferences of a certain voter. In our case, the set of rankings based on corresponding bibliometric indicators is treated as a profile of opinions of seven virtual experts.

Let A denote the set of feasible alternatives; let N denote a group of experts making a collective decision by vote. A decision is a choice of a subset from A . *Preferences* of a voter i , $i \in N$, are revealed through pairwise comparisons of alternatives and are modeled by a binary relation P_i on A , $P_i \subseteq A \times A$: if voter i prefers x to y , then the ordered pair (x, y) belongs to the relation P_i . If a voter is unable to compare two alternatives or thinks they are of equal value, it will be presumed that he is indifferent regarding the choice between them.

If chooser's preferences are known and a *choice rule* (a mapping of the set of binary relations on A onto the set of nonempty subsets of A) is given, then it is possible to determine what alternatives should be the result of her choice. Thus a social choice problem can be solved if one knows voters' preferences (experts' opinions), defines a binary relation μ , $\mu \subseteq A \times A$, that models social preferences (group's opinion), and determines a social choice rule $S(\mu, A): \{\mu\} \rightarrow 2^A \setminus \emptyset$. Probably the most popular method to construct μ is to apply *the majority rule*: (x, y) belongs to μ if the number of those who think x is better than y is greater than the number of those who think y is better than x : $x\mu y \Leftrightarrow |N_1| > |N_2|$, where $N_1 = \{i \in N \mid xP_i y\}$, $N_2 = \{i \in N \mid yP_i x\}$. In this case, μ is called *the majority relation*.

The choice of this particular rule of aggregation is prescribed by the social choice theory since the majority rule, and this rule only, satisfies several important normative conditions (May 1952), such as *independence of irrelevant alternatives*, *Pareto-efficiency*, *neutrality* (equal treatment of alternatives), and *anonymity* (equal treatment of voters).

The majority relation quite often happens not to be a ranking itself since it is generally nontransitive. That is, the majority relation may contain cycles. This result is known as *the*

Condorcet paradox (Condorcet, 1785). In order to check if the majority relation is transitive or not and to evaluate how nontransitive it is, we calculate the number of 3-step μ -cycles, 4-step μ -cycles and 5-step μ -cycles for three sets of journals (Table 2).

Table 2. Numbers of 3-, 4- and 5-step μ -cycles for three sets of journals.

	<i>3-step cycles</i>	<i>4-step cycles</i>	<i>5-step cycles</i>
Economics	2446	22427	226103
Management	203	787	3254
Political Science	149	430	1344

As we see, the Condorcet paradox occurs in all three cases.

The Copeland rule

In order to bypass the nontransitivity problem, several methods have been proposed. Probably, the simplest one is the Copeland rule (Copeland, 1951). The idea behind it is the following: the greater the number of alternatives that are worse than a given one, the better this alternative is (the 2nd version of the Copeland rule); and it is determined through pairwise comparisons whether a given alternative is either better or worse than another one. Alternatively, it could be put that an alternative is good if the number of alternatives that are better is small (the 3rd version of the rule). Finally, one can subtract the number of alternatives that are more (socially) preferable than a given one from the number of alternatives less preferable and then rank alternatives by values of these differences (1st version of the rule). All three versions yield the same result when there are no ties. We used the second and the third versions of the Copeland rule.

A sorting procedure based on tournament solutions

In order to construct a ranking, we can use solutions to the problem of optimal social choice. A solution concept $S(\mu, A)$ is a choice rule that determines a set $B_{(1)}$ of those alternatives that are considered to be social optima: $B_{(1)}=S(\mu, A)$. Let us exclude them and repeat the sorting procedure for the subset $A \setminus B_{(1)}$. The set $B_{(2)}=S(\mu, A \setminus B_{(1)})=S(\mu, A \setminus S(\mu, A))$ contains second best choices, for they are worse than alternatives from $B_{(1)}$ and better than options from $A \setminus (B_{(1)} \cup B_{(2)})$. After a finite number of selections and exclusions, all alternatives from A will be separated by classes $B_{(k)}=S(\mu, A \setminus (B_{(k-1)} \cup B_{(k-2)} \cup \dots \cup B_{(2)} \cup B_{(1)}))$ according to their “quality”, and these classes constitute a ranking.

We use two choice rules called *tournament solutions*: the uncovered set (Miller, 1980) and the externally stable set (von Neumann & Morgenstern, 1944; Aleskerov & Kurbanov 1999; Subochev, 2008; Aleskerov & Subochev, 2013). The former is based on the idea of choosing “strong” candidates; the latter chooses candidates from “strong” groups.

We say that an alternative x *covers* (meaning that it is definitively better than) an alternative y if x is (socially) preferred not only to y but also to all alternatives that are less preferable than y : $x\mu y \wedge \forall z \in A, y\mu z \Rightarrow x\mu z$. The *uncovered set* UC is comprised of all alternatives that are not covered by any other alternative.

The concept of a minimal externally stable set operationalizes the idea of a strong group of candidates. A set ES is *externally stable* if for any alternative x outside ES there exists an

This work is licensed under a Creative Commons License: Attribution-NonCommercial-NoDerivatives 4.0 International.

alternative y in ES that is more preferable (socially) than x : $\forall x \notin ES, \exists y: y \in ES \wedge y \mu x$. An externally stable set is *minimal* if none of its proper subsets is externally stable. An alternative is regarded to be optimal if it belongs to some minimal externally stable set; therefore, the solution is the union of all such sets and is denoted MES .

Both UC and MES are always nonempty and can be calculated through their matrix-vector representations given by Aleskerov and Subochev (2013).

The Markovian method

Finally, we apply a version of a ranking procedure called the Markovian method since it is based on an analysis of Markov chains that model stochastic moves from vertex to vertex via arcs of a digraph representing a binary relation μ . The earliest versions of this procedure were proposed by Daniels (1969) and Ushakov (1971). A similar method has been introduced in bibliometrics by Pinsky and Narin (1976). The detailed description of the procedure is given in (Aleskerov, Pislyakov & Subochev 2014).

The table with ranks of all journals in all rankings can be found in (Aleskerov, Pislyakov & Subochev 2014) as well.

CORRELATION ANALYSIS

To evaluate the (in)consistency of two rankings, we measure their correlation. In this paper, we use the Kendall rank correlation coefficient τ_b . Table 3 visualizes its values for all pairs of rankings, initial and aggregate. The corresponding numerical values of τ_b can be found in (Aleskerov, Pislyakov & Subochev 2014).

In all cases, ranking by values of the immediacy index demonstrates the lowest level of correlation with single-indicator-based rankings. This is possibly due to a very narrow publication window that this indicator is based on. In all cases, rankings based on the 5-year impact factor demonstrate the highest level of correlation among single-indicator-based rankings. In the previous study (Aleskerov et al. 2011), the most correlated ranking was one based on the classic impact factor, the 5-year impact being the second best. Systematic differences between rankings based on other indicators are not observed.

Direct observations of values of τ_b for pairs with an aggregate ranking confirm our previous results (Aleskerov et al. 2011). For each set of journals, all aggregate rankings correlate with any single-indicator-based ranking better than other single-indicator-based rankings do. The only exception is correlation of impact factor with 5-year IF, which is a bit higher than correlation with aggregate rankings. This is not true for 5-year IF, though. Formal comparisons based on majority rule (see (Aleskerov, Pislyakov & Subochev 2014) for details) confirm direct observations. In all cases, almost all aggregate ranking methods produce rankings that represent the set of single-indicator-based rankings better than any of these seven. Therefore replacing the set of seven single-indicator-based rankings with aggregate rankings is justified, the best method producing the most representative rankings being the third version of the Copeland rule.

CONCLUSION

Measuring journal influence is a problem that has no clear-cut solution. Different approaches lead to different indicators, and each possesses its own justification. We took the values of seven popular bibliometric indicators as our data. The correlation analysis has shown that the 5-year impact factor is the best choice if one tries to represent seven single-indicator-based journal rankings by one of them. The least correlated are rankings based on the immediacy index. Other indicators are of more or less equal representativeness.

Despite the correlation of single-indicator-based rankings being high, there is a significant number of contradictions. We propose to minimize their number by replacing the set of rankings with an aggregate ranking. Aggregation can be performed in many ways. This report demonstrates the power of ordinal methods borrowed from social choice theory. This is a novel approach in bibliometrics. Ordinal procedures relieve a researcher from the burden of finding appropriate weights and theoretical justifications for arithmetic operations with aggregated variables. The correlation analysis has also shown that aggregate rankings reduce the number of contradictions and represent the set of single-indicator-based rankings better than any of the seven rankings themselves. Thus, aggregate rankings are more efficient instruments for the evaluation of journal influence.

Table 3. Kendall rank correlation coefficient τ_b visualized through a greyscale (the higher is the value, the darker is the cell; pure white corresponds to $\tau_b < 0.5$, pure black – to $\tau_b > 0.95$, the scale interval is 0.05).

	<i>IF</i>	<i>5-year IF</i>	<i>immediacy index</i>	<i>article influence</i>	<i>h-index</i>	<i>SNIP</i>	<i>SJR</i>	<i>Copeland (2)</i>	<i>Copeland (3)</i>	<i>UC</i>	<i>MES</i>	<i>Markovian</i>
Economics												
<i>IF</i>	Black	Dark	White	Light	Light	Light	Light	Light	Light	Light	Light	Light
<i>5-year IF</i>	Dark	Black	White	Light	Light	Light	Light	Light	Light	Light	Light	Light
<i>immediacy index</i>	White	White	Black	Light	Light	Light	Light	Light	Light	Light	Light	Light
<i>article influence</i>	Light	Light	White	Black	Light	Light	Light	Light	Light	Light	Light	Light
<i>h-index</i>	Light	Light	Light	Light	Black	Light	Light	Light	Light	Light	Light	Light
<i>SNIP</i>	Light	Light	Light	Light	Light	Black	Light	Light	Light	Light	Light	Light
<i>SJR</i>	Light	Light	Light	Light	Light	Light	Black	Light	Light	Light	Light	Light
Management												
<i>IF</i>	Black	Dark	White	Light	Light	Light	Light	Light	Light	Light	Light	Light
<i>5-year IF</i>	Dark	Black	White	Light	Light	Light	Light	Light	Light	Light	Light	Light
<i>immediacy index</i>	White	White	Black	Light	Light	Light	Light	Light	Light	Light	Light	Light
<i>article influence</i>	Light	Light	White	Black	Light	Light	Light	Light	Light	Light	Light	Light
<i>h-index</i>	Light	Light	Light	Light	Black	Light	Light	Light	Light	Light	Light	Light
<i>SNIP</i>	Light	Light	Light	Light	Light	Black	Light	Light	Light	Light	Light	Light
<i>SJR</i>	Light	Light	Light	Light	Light	Light	Black	Light	Light	Light	Light	Light
Political Science												
<i>IF</i>	Black	Dark	White	Light	Light	Light	Light	Light	Light	Light	Light	Light
<i>5-year IF</i>	Dark	Black	White	Light	Light	Light	Light	Light	Light	Light	Light	Light
<i>immediacy index</i>	White	White	Black	Light	Light	Light	Light	Light	Light	Light	Light	Light
<i>article influence</i>	Light	Light	White	Black	Light	Light	Light	Light	Light	Light	Light	Light
<i>h-index</i>	Light	Light	Light	Light	Black	Light	Light	Light	Light	Light	Light	Light
<i>SNIP</i>	Light	Light	Light	Light	Light	Black	Light	Light	Light	Light	Light	Light
<i>SJR</i>	Light	Light	Light	Light	Light	Light	Black	Light	Light	Light	Light	Light

Some of the aggregate rankings (produced by the Copeland rule and the Markovian method) are characterized by a high level of discrimination, and their shares of tied pairs are very small (less than 1%). For instance, the Markovian method discriminate almost all journals. Other rankings (those based on tournament solutions) are rough orderings, which could also be of

value. One may even argue that these rough orderings, when many journals are regarded as equal to each other, better represent our intuitive judgments concerning journal influence.

Not all social choice ranking methods have been employed in this study. There are also other tournament solutions. The next logical step would be to widen both the arsenal of aggregation techniques and the set of empirical data.

REFERENCES

- Aleskerov, F. & Kurbanov, E. (1999). Degree of manipulability of social choice procedures. In A. Alkan, Ch.D. Aliprantis N.C. Yannelis (Eds.), *Current Trends in Economics: Theory and Applications* (pp. 13–27). N.Y.: Springer-Verlag.
- Aleskerov, F.T., Pislyakov, V.V. & Subochev, A.N. (2014). *Ranking Journals in Economics, Management and Political Science by Social Choice Theory Methods*. WP BRP 27/STI/2014. Moscow: HSE.
- Aleskerov, F.T., Pislyakov, V.V., Subochev, A.N. & Chistyakov, A.G. (2011). *Rankings of Management Science Journals Constructed by Methods from Social Choice Theory*. Working paper WP7/2011/04. Moscow: Higher School of Economics. (in Russian).
- Aleskerov, F. & Subochev, A. (2013). Modeling optimal social choice: Matrix-vector representation of various solution concepts based on majority rule. *Journal of Global Optimization*, 56(2), 737–756.
- Arrow, K.J. (1951). *Social Choice and Individual Values*. New York: Wiley.
- Arrow, K.J. & Raynaud, H. (1986). *Social Choice and Multicriterion Decision-Making*. Cambridge (Mass.): MIT Press.
- Condorcet, Marquis de. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Paris: L'imprimerie royale.
- Copeland, A.H. (1951). *A reasonable social welfare function*, Seminar on Application of Mathematics to the Social Sciences, University of Michigan, Ann Arbor. Mimeo.
- Daniels, H.E. (1969). Round-robin tournament scores. *Biometrika*, 56(2), 295–299.
- Garfield, E. & Sher, I.H. (1963). New Factors in the Evaluation of Scientific Literature through Citation Indexing. *American Documentation*, 14(3), 195–201.
- Gross, P.L.K. & Gross, E.M. (1927). College libraries and chemical education. *Science*, 66(1713), 385–389.
- Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46), 16569–16572.
- Miller, N.R. (1980). A new solution set for tournaments and majority voting: Further graph-theoretical approaches to the theory of voting. *American Journal of Political Science*, 24(1), 68–96.
- May, K.O. (1952). A set of independent necessary and sufficient conditions for simple majority decisions. *Econometrica*, 20(4), 680–684.
- von Neumann, J. & Morgenstern O. (1944). *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.
- Pinski, G. & Narin, F. (1976). Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics. *Information Processing and Management*, 12(5), 297–312.
- Pislyakov, V. (2009). Comparing two “thermometers”: Impact factors of 20 leading economic journals according to Journal Citation Reports and Scopus. *Scientometrics*, 79(3), 541–550.
- Subochev, A. (2008). *Dominant, Weakly Stable, Uncovered Sets: Properties and Extensions*. Working paper WP7/2008/03. Moscow: SU – Higher School of Economics.
- Ushakov, I.A. (1971). The problem of choosing the preferable object. *Izvestiya Akademii Nauk SSSR. Tekhnicheskaya Kibernetika*, 4, 3–7. (in Russian)



The performance and trend of China's academic disciplines from 2006 to 2014¹

Hu Zhigang*, Huang Fu**, Liang Guoqiang**, Wang Xianwen** and Tang Delong**

*huzhigang@dlut.edu.cn

WISE Lab, Dalian University of Technology, Dalian, 116024 (P.R. China)

DUT – Drexel Joint Institute for the Study of Knowledge Visualization and Scientific Discovery, Dalian University of Technology, Dalian 116085, China

** hf206@163.com; liang_1988@mail.dlut.edu.cn; xianwenwang@dlut.edu.cn; tangdl@mail.dlut.edu.cn
WISE Lab, Dalian University of Technology, Dalian, 116024 (P.R. China)

ABSTRACT

China has achieved remarkable growth in science and technology production. But different academic disciplines vary greatly in size and performance. In this paper, the performance of each China's academic discipline was measured by counting its scientific outputs in national and international literature databases: CNKI and WOS. The results show that China's preferential research areas in national journals (CNKI) is different from those in international journals (WOS). On time dimension, some research areas (i.e. Environmental Science and Engineering) are getting hotter and others (i.e. Computer Science) go the opposite way.

INTRODUCTION

In the past decade, China has achieved remarkable growth in science and technology production. In 2006, China published about 90,000 papers indexed by SCI, which accounted for 7.1% globally and ranked fifth. While in 2015, the SCI publications has increased to almost 300,000 (16.6%), only behind the United States. But if looking closely, you can see that different academic disciplines vary greatly in size and performance. China has already be a leading nation in hard science, such as engineering, energy, materials science, and computer science, but in the soft science like psychology, arts & humanities, and social science, China is still far behind the world-class scientific power(Wang, 2016).

Many papers have studied the performances of China's academic disciplines. The performance of a academic discipline is generally measured by the amount of scientific publications (Guan & Gao, 2008; Guan & Ma, 2004a; Guan & Wang, 2010; Liu, Xu, & Li, 2015) or their citation times (Kostoff, 2008; Moiwo & Tao, 2012; Yang, Ma, Song, & Qiu, 2010). For example, Wang L. (Wang, 2016) evaluated China's research performance of each discipline by using its scientific output in journals sourced from Elsevier's Scopus. Guan et al. conducted a serial of researches with Web of Science to examine China's performance in some specific disciplines (Gao & Guan, 2009; Guan & Gao, 2008; Guan & He, 2005; Guan & Ma, 2004b, 2007; Guan & Wang, 2010). Among these researches, Web of Science, Scopus, Engineering Village (EI) and other international literature databases are the most common data sources. However, international journals are not the only, even the primary choice for Chinese scholars. In China, most of research papers are written in Chinese, published in Chinese journals, and read, obviously, by Chinese scholars. This is especially true for those

¹ This work was supported by National Natural Science Foundation of China (NSFC 71503031)

disciplines like social sciences and humanities. In these fields, research topics are usually too national to raise the interest from abroad.

So far, there are few researches on China's national productivity, though it's actually more sufficient indicator to reveal China's performance in each discipline (Zhou & Leydesdorff, 2006). In 2002, Moed assessed China's research activities by distinguish between a national and an international point of view using Chinese Science Citation Database (CSCD) and ISI Indexes (Moed, 2002). By comparing China's national and international outputs, he got two rather conflicted pictures about China's preponderant disciplines. In this paper, we would compare China's international and national performances again with data from WOS and CNKI, a much better coverage of Chinese periodicals than CSCD. It is still a significant issue for getting an in-depth knowledge of China's discipline and their internationalize level.

DATA COLLECTION

In this paper, the performance of each China's academic discipline was measured by counting its scientific outputs in national and international literature databases. For its international performance, Web of Science (SCI, SSCI and A&HCI), the most popular citation database in the world, was chosen as data source. For its national performance, China Academic Journal Database of CNKI[®], one of the largest scientific literature databases in China was chosen respondingly.

Web of Science

In the newest (version 5) Web of Science, including natural sciences(SCI), social sciences(SSCI), and arts and humanities(AHCI), 127,00 journals are divided into 256 WOS Categories (WC) or 151 Research Areas (SC). The difference between WCs and SCs is that the designation of WCs is journal-based while SCs is article-based. However, we didn't find any description on how an article's SC is determined. And as we find, the designation of SCs is still journal-based basically: almost all the articles in the same journal will be assigned to the same SC. In this research, the set of SC, which is at a higher level of aggregation than WC, is chosen as classification system.

CNKI

The academic journal database of CNKI is the largest Chinese journal database. It collected 10,116 China academic journals, covering the areas of science, engineering technology, agriculture, philosophy, medicine, humanities and social sciences, etc. Now the full-text paper amount in the database has reached to 51 million. In CNKI, the articles is assigned into 10 collections, i.e. Science/Technology/Engineering I, II & III, Agriculture, Medicine/Hygiene, Literature/History/Philosophy, Politics/Military/Law, Education/Social Science, Electronics/Information Technology, Economics and Management. The 10 collections are further divided into 168 Research Subjects (RS).

The classification of RS in CNKI is article-based. An article is assigned to one more RSEs by analysing its title and keywords and referring to Classified Chinese Thesaurus, a compiled integrated and classified subject thesaurus developed with Chinese Library Classification (CLC). Compared with WOS, the category in CNKI is more fine-grained and the classification method is more rigorous in the way determining the RSEs of articles.

RESULTS

Performance of China's research disciplines in WOS

For each research discipline, its publication and proportion is computed and listed in Table 1. In Web of Science, China's academic research focuses primarily on the area of physical Sciences such as "Materials Science", "Chemistry", and "Physics Applied". In 2014, for example, Chinese scholars published more than 15,000 papers in each of these three research areas. They contributed 8.99%, 7.07% and 5.70% of China's total international output respectively, and the shares were still increasing.

Table 1. The top 10 most productivity disciplines in WOS

WOS Category	2006	2008	2010	2012	2014
Materials Science Multidisciplinary	8369	11498	13384	17845	24445
	7.10%	9.75%	8.94%	8.94%	8.99%
Chemistry Multidisciplinary	6803	7292	9505	13208	19213
	5.77%	6.19%	6.35%	6.62%	7.07%
Physics Applied	5535	7609	8471	11496	15484
	4.69%	6.45%	5.66%	5.76%	5.70%
Chemistry Physical	5531	8261	9641	11997	14987
	4.69%	7.01%	6.44%	6.01%	5.51%
Engineering Electrical Electronic	3270	4613	6538	9067	12909
	2.77%	3.91%	4.37%	4.54%	4.75%
Biochemistry Molecular Biology	3859	5077	6347	8718	10957
	3.27%	4.31%	4.24%	4.37%	4.03%
Oncology	1197	1930	3210	5714	10952
	1.02%	1.64%	2.14%	2.86%	4.03%
Nanoscience Nanotechnology	1707	3282	4395	6136	9321
	1.45%	2.78%	2.93%	3.07%	3.43%
Optics	2411	3773	5051	6425	8463
	2.05%	3.20%	3.37%	3.22%	3.11%
Energy Fuels	912	1750	2918	4296	8009
	0.77%	1.48%	1.95%	2.15%	2.95%

To identify the trends of China's disciplines since 2006, Table 2 lists the top 10 fastest-growing. During the past ten years, "Medicine Research Experimental" is the China's fastest growing research area. Its share increases from 0.49% to 2.10%, and its ranking rises from 66 to 21 among all the WOS subjects. "Oncology", another discipline of medicine sciences, grown from 1.02% and #35 in 2006 to 4.03% and #8 in 2014. There are also two research areas belonging to medicine sciences in the top ten list, "Cardiac Cardiovascular Systems" (from #73 to #41) and "Surgery" (from #60 to #40). Besides, the research areas of "Energy Fuels", "Nanoscience and Nanotechnology", "Engineering Environmental" and "Thermodynamics" also perform very well in the past ten years.

Table 2. The fastest-growing and fastest-falling disciplines in WOS

WOS Category	2006		2008		2010		2012		2014		Growth
Medicine Research Experimental	0.49%	66	0.92%	51	1.05%	47	1.43%	33	2.10%	21	324.49%
Oncology	1.02%	35	1.64%	26	2.14%	20	2.86%	11	4.03%	8	296.95%
Energy Fuels	0.77%	49	1.48%	30	1.95%	24	2.15%	21	2.95%	11	281.11%
Food Science Technology	0.49%	67	0.92%	52	1.21%	39	1.39%	34	1.40%	34	186.71%
Cardiac Cardiovascular Systems	0.45%	73	0.68%	68	1.96%	23	1.16%	42	1.26%	41	182.77%
Nanoscience Nanotechnology	1.45%	24	2.78%	12	2.93%	12	3.07%	9	3.43%	9	136.81%
Engineering Environmental	0.52%	63	0.90%	55	1.23%	38	1.19%	40	1.07%	49	105.16%
Surgery	0.62%	60	0.99%	45	1.24%	37	1.28%	37	1.28%	40	105.14%
Thermodynamics	0.54%	62	0.76%	64	0.84%	64	0.77%	66	1.10%	47	103.53%
Electrochemistry	0.91%	39	1.49%	29	1.55%	31	1.76%	26	1.83%	25	101.10%
Computer Science Artificial Intelligence	<u>2.92%</u>	<u>9</u>	<u>1.02%</u>	<u>44</u>	<u>1.16%</u>	<u>43</u>	<u>1.18%</u>	<u>41</u>	<u>1.23%</u>	<u>43</u>	<u>-57.96%</u>
Physics Multidisciplinary	3.87%	5	5.05%	5	3.86%	7	3.29%	7	2.30%	18	-40.48%
Metallurgy Metallurgical Engineering	3.59%	6	4.01%	7	3.23%	11	2.66%	16	2.28%	20	-36.44%
Chemistry Inorganic Nuclear	1.81%	18	2.10%	22	1.93%	25	1.70%	27	1.39%	36	-23.12%
Polymer Science	2.56%	15	2.78%	13	2.52%	16	2.34%	19	1.98%	22	-22.79%

The top 5 most-decreasing research areas is also listed in Table 2. Surprisingly, “Computer Science and Artificial Intelligence”, one of the most hot academic disciplines in China, has decreased by 57.96% since 2006, and its ranking dropped from #9 to #43. Some other previously dominant discipline, like “Physics Multidisciplinary”, “Metallurgy Metallurgical Engineering” and “Chemistry Inorganic Nuclear”, all experienced a serious drop in shares and a sharply falling in ranking this years.

Performance of China’s research disciplines in CNKI

In 2014, the most productivity research area is “Environment Science and Resources Utilization”, which accounts for 3.21% of all publications. “Light Industry, Handicraft Industry” ranked second with 3.04%. The third-ranking is the field of “Electric Power Industry”. Besides, the proportions of “Architecture and Engineering”, “Metal Science and Metal Technics” and “Higher Education” were also quite high, as it shown in Table 3.

Table 3. The top 10 most productivity disciplines in CNKI

CNKI Category	2006	2008	2010	2012	2014
Environment Science and Resources Utilization	11942	15454	16748	17287	18603
	2.27%	2.62%	2.69%	3.05%	3.21%
Light Industry, Handicraft Industry	13177	14073	15392	17323	17616
	2.50%	2.39%	2.47%	3.06%	3.04%
Electric Power Industry	12371	13817	15286	15242	16769
	2.35%	2.34%	2.46%	2.69%	2.90%
Architecture and Engineering	12820	14598	16709	15620	16684
	2.44%	2.48%	2.69%	2.76%	2.88%
Macro-economic Management and Sustainable Development	14811	17312	18289	22325	16612
	2.82%	2.94%	2.94%	3.94%	2.87%
Higher Education	9095	12567	13104	13802	15096
	1.73%	2.13%	2.11%	2.44%	2.61%
Computer Software and Application of Computer	20203	20987	18459	14221	14800
	3.84%	3.56%	2.97%	2.51%	2.56%
Metal Science and Metal Technics	9687	11999	13228	13195	14402
	1.84%	2.03%	2.13%	2.33%	2.49%
Chemistry	8899	7848	7465	10801	13567
	1.69%	1.33%	1.20%	1.91%	2.34%
Organic Chemical Industry	12953	13821	15262	13777	13128
	2.46%	2.34%	2.45%	2.43%	2.27%

Table 4 lists the most changed academic disciplines from 2006 to 2014 in CNKI. “Security Science and Disaster Prevention”, “General Chemistry Industry” and “Fundamental Science of Agriculture” are the top three fastest rising academic disciplines. “Mathematics”, “Internet Technology”, “Industrial Current Technology and Equipment” are the top three fastest falling ones. Impressively, the ranking of “Computer Software and Application of Computer” drop from the first in 2006 to the seventh in 2014; while “Environment Science and Resources Utilization” rose from the tenth to the first conversely.

Generally speaking, significant rise also happened in the academic disciplines related to resource and environment, such as the fields of “Security Science and Disaster Prevention”, “Mining Engineering”, “Meteorology”. The medicine sciences, however, have experienced a clear falling since 2006. For example, “Ophthalmology and Otolaryngology”, “Urology”, “Traditional Chinese Medicine”, “Surgery”, “Fundamental Medicine” have all dropped by 30% or more.

Table 4. The top 10 fastest-growing and fastest-falling disciplines in CNKI

CNKI Research Area	2006		2008		2010		2012		2014		Growth
Security Science and Disaster Prevention	0.34%	109	0.54%	79	0.55%	81	0.65%	70	0.63%	66	84.18%
General Chemistry Industry	0.28%	123	0.33%	107	0.37%	99	0.51%	83	0.50%	83	76.23%
Fundamental Science of Agriculture	0.68%	65	0.87%	51	0.98%	40	0.94%	43	1.14%	33	68.07%
Mining Engineering	1.15%	37	1.44%	26	1.75%	18	1.89%	17	1.92%	16	66.47%
Meteorology	0.35%	105	0.48%	86	0.55%	79	0.48%	88	0.55%	77	55.85%
Physical Geography and Topography	0.42%	91	0.53%	80	0.54%	82	0.49%	86	0.63%	65	49.76%
Railway Transportation	0.43%	89	0.47%	88	0.47%	86	0.52%	82	0.61%	68	42.24%
Environment Science and Resources Utilization	<u>2.27%</u>	<u>10</u>	<u>2.62%</u>	<u>4</u>	<u>2.69%</u>	<u>4</u>	<u>3.05%</u>	<u>3</u>	<u>3.21%</u>	<u>1</u>	<u>41.57%</u>
Animal Husbandry and Veterinary	1.21%	32	1.55%	23	1.69%	21	1.76%	21	1.65%	22	35.91%
Metal Science and Metal Technics	1.84%	16	2.03%	14	2.13%	12	2.33%	9	2.49%	8	35.11%
Mathematics	2.42%	8	1.62%	20	1.50%	25	1.59%	25	0.95%	42	-60.58%
Internet Technology	0.82%	50	0.61%	72	0.48%	84	0.43%	92	0.37%	102	-55.07%
Industrial Current Technology and Equipment	1.06%	40	1.29%	33	1.11%	37	0.90%	48	0.59%	72	-44.67%
Ophthalmology and Otolaryngology	0.76%	55	0.52%	82	0.51%	83	0.41%	94	0.43%	92	-43.82%
Biomedicine Engineering	0.36%	100	0.41%	93	0.30%	113	0.20%	135	0.20%	134	-42.96%
Urology	0.57%	76	0.50%	85	0.45%	88	0.35%	108	0.33%	113	-41.20%
Traditional Chinese Medicine	1.36%	25	0.89%	49	0.95%	41	0.78%	57	0.81%	55	-40.61%
Surgery	2.00%	14	1.59%	21	1.54%	23	1.50%	26	1.30%	27	-35.16%
Fundamental Medicine	1.23%	30	0.94%	43	0.86%	49	0.90%	49	0.81%	54	-33.73%
Computer Software and Application of Computer	<u>3.84%</u>	<u>1</u>	<u>3.56%</u>	<u>1</u>	<u>2.97%</u>	<u>1</u>	<u>2.51%</u>	<u>6</u>	<u>2.56%</u>	<u>7</u>	<u>-33.42%</u>

DISCUSSION

The results showed China's research disciplinary performance and trend both in international and national journals. The trend of each research area since 2006 is revealed. Discussion will focus on the following questions: (a) why China's preferential research areas in national journals (CNKI) is different from those in international journals (WOS)? and (b) why some research areas are getting hotter and others go the opposite way?

Why China's hot disciplines in CNKI is different from WOS?

Obviously, the hot research areas in national journals (CNKI) are different with those in WOS. There is two reasons related to this. Firstly, WOS and CNKI adopted totally different Subject Classification. Most research areas in WOS is hard to find the completely same ones in CNKI. CNKI category, an typical China subject category, looks more coarse-grained in size, and more out-dated compared with the category in WOS. For example, "Physical Chemistry" is an individual discipline in WOS; while an sub-discipline of "Chemistry" in CNKI. "Optics" is separated in WOS and a branch of the discipline of "Physics" in CNKI. Similarly, "Nanoscience Nanotechnology", a new rising research area, is already an

independent academic discipline in WOS, the same as “Materials Science” and “Chemistry”; while in CNKI “Nano Materials” is only a branch of “Special Structured Material”, a sub-discipline under the discipline of “Materials Science”. Table 7 shows the hot disciplines in WOS and their responding academic disciplines in CNKI that we could find. You can see from this table the difference between these two categories.

Secondly, some research areas are more internationalized than others. It is easy to understand that social sciences are usually less inclined to publish papers in foreign journals. Actually, even for the natural science, different subjects’ level of internationalization varied greatly.

Why some research areas are getting hotter and other go the opposite way?

The state of economic of China is the main reason of the evolution of China’s academic researches. As we know, China’s economic growth is relying heavily on processing, manufacturing and infrastructure construction. Especially in 2008, to cope with the international financial crisis and severe economic recession, China government responded by announcing a 4 trillion Yuan (\$586 billion) spending package, driven by massive infrastructure investments across the country, and, from then on, the research areas related with construction industry, for example, “Architecture and Engineering”, “Electric Power Industry”, and “Mining Engineering”, are starting to heat up. Besides that, the increase of “Light Industry, Handicraft Industry” proves China’s emergence as a global manufacturing power.

The social concerning is another important factors in depending which areas is getting hotter. Recently, China begins to give great emphasis to environmental concerns, so the sharp rise has occurred in the fields of “Environment Science and Resources Utilization”. Moreover, following by the understanding and implementation of China government’s strategy of invigorating the country through science and education, the development of “higher education” has realized substantial growth.

In addition to the rising research fields we’ve discussed, we also find some areas is decreasing. The proportion of “Computer Software and Application of Computer” has dropped by 33.42% (CNKI) and 57.96% (WOS) since 2006. This is because each subject has its own life cycle. Generally, when a research area has began to be industrial and technological, the theoretical research will fall and fade. That’s exactly what happened in the field of computer science.

REFERENCES

- Gao, X., & Guan, J. (2009). A scale-independent analysis of the performance of the Chinese innovation system. *Science And Technology*, 3, 321–331. doi:10.1016/j.joi.2009.04.004
- Guan, J., & Gao, X. (2008). Comparison and evaluation of Chinese research performance in the field of bioinformatics. *Scientometrics*, 75(2), 357–379. doi:10.1007/s11192-007-1871-0
- Guan, J., & He, Y. (2005). Comparison and evaluation of domestic and international outputs in Information Science & Technology research of China. *Scientometrics*, 65(2), 215–244. doi:10.1007/s11192-005-0269-0

- Guan, J., & Ma, N. (2004a). A comparative study of research performance in computer science. *Scientometrics*, 61(3), 339–359. doi:10.1023/B:SCIE.0000045114.85737.1b
- Guan, J., & Ma, N. (2007). A bibliometric study of China's semiconductor literature compared with other major Asian countries. *Scientometrics*, 70(1), 107–124. doi:10.1007/s11192-007-0107-7
- Guan, J., & Wang, G. (2010). A comparative study of research performance in nanotechnology for China's inventor–authors and their non-inventing peers. *Scientometrics*, 84(2), 331–343. doi:10.1007/s11192-009-0140-9
- Kostoff, R. (2008). Comparison of China/USA science and technology performance ☆. *Journal of Informetrics*, 2(4), 354 – 363. doi:10.1016/j.joi.2008.06.004
- Liu, X., Xu, Q., & Li, M. (2015). A comparative analysis of scientific publications in management journals by authors from Mainland China, Hong Kong, Taiwan, and Macau: 2003–2012. *Scientometrics*, 105(1), 135–143. doi:10.1007/s11192-015-1665-8
- Moed, H. F. (2002). Measuring China's research performance using the Science Citation Index. *Scientometrics*, 53(3), 281–296.
- Moiwo, J. P., & Tao, F. (2012). The changing dynamics in citation index publication position China in a race with the USA for global leadership. *Scientometrics*, 95(3), 1031–1050. doi:10.1007/s11192-012-0846-y
- Wang, L. (2016). The structure and comparative advantages of China's scientific research: quantitative and qualitative perspectives. *SCIENTOMETRICS*, 106(1), 435–452. doi:10.1007/s11192-015-1650-2
- Yang, S., Ma, F., Song, Y., & Qiu, J. (2010). A longitudinal analysis of citation distribution breadth for Chinese scholars. *Scientometrics*, 85(3), 755–765. doi:10.1007/s11192-010-0245-1
- Zhou, P., & Leydesdorff, L. (2006). The Emergence of China as a Leading Nation in Science. *Research Policy*, 35(1), 83–104.



Comparing absolute and normalized indicators in scientific collaboration: a study in Environmental Science in Latin America

Maria Cláudia Cabrini Grácio* and Ely Francina Tannuri de Oliveira**

*cabrini@marilia.unesp.br

UNESP – Univ Estadual Paulista, 737 Hygino Muzzi Filho Avenue, 17525-900 Marília (Brazil)

**etannuri@gmail.com

UNESP – Univ Estadual Paulista, 737 Hygino Muzzi Filho Avenue, 17525-900 Marília (Brazil)

ABSTRACT

This paper aims to conduct a comparative analysis of scientific collaboration proximity trends generated from absolute indicators and indicators of collaboration intensity in the field of Environmental Sciences in Latin America (LA), in order to identify possible existing biases in the absolute indicators of international cooperation, due to the magnitude of scientific production of these countries in mainstream science.

More specifically, the objective is to analyze the compared forms of absolute and normalized values of co-authorship among Latin America countries and their main collaborators, in order to observe similarities and differences expressed by two indexes of frequency in relation to scientific collaboration trends in LA countries. In addition, we aim to visualize and analyze scientific collaboration networks with absolute and normalized indexes of co-authorship through SC among Latin America countries and their collaborators, comparing proximity evidenced by two generated collaborative networks - absolute and relative indicators.

Data collection comprised a period of 10 years (2006-2015) for the countries from LA: Brazil, Mexico, Argentina, Chile and Colombia as they produced 94% of total production, a percentage considered representative and significant for this study.

Then, we verified the co-authorship frequencies among the five countries and their key collaborators and built the matrix with the indexes of co-authorship normalized through SC. Then, we generated two egocentric networks of scientific collaboration - absolute frequencies and normalized frequencies through SC using Pajek software.

From the results, we observed the need for absolute and normalized indicators to describe the scientific collaboration phenomenon in a more thoroughly way, once these indicators provide complementary information.

INTRODUCTION

Latin America (LA) comprises a territory of 20 countries at different levels of social, economic and political development.

After World War II, between 1960s and 1970s, there was an increase in the creation of institutions responsible for promoting policies and instruments to guide and manage the scientific and technological development of LA countries (Velho, 2005).

In the 80s, a regional crisis responsible for a period of reduced economic resources with political, economic and social changes, resulted in the discontinuance of efforts made in previous years in the field of higher education in these countries.

In the 90s, policies have sought mainly to develop an openness of the economy, macroeconomic stability and competitiveness in international markets. These factors were responsible for the changes that subsequently elapsed in Science and Technology systems in these countries (Gazzola & Didriksson, 2008; Velho, 2005).

Regarding publications in mainstream science, according to SCImago Journal & Country Rank (JCR), Latin America is one of the regions that grew the most in relation to the world scientific production in recent years (SCImago Research Group, 2015). In the period studied in this research, the total of international scientific production increased 178%, while LA reached 477%, continuously, above the global total, with an average of 8.9 citations per paper (Gutierrez & Gracio, 2015).

In the period 2003-2012, Latin America published in Web of Science (WoS) more than 730,000 documents and in Scopus, more than 880,000 documents. According to these data, it is noted that the total scientific production in South America has grown from about 2% of world production in 1996 to 4% in 2012, and currently represents about 3% of total publications, both in WoS and in Scopus, presenting higher growth in comparison with the global growth in the period (32% versus 19%) (Gutierrez & Gracio, 2015).

As a result of the significant increase in Latin American participation in mainstream science, observed in international databases, the use of methodologies to analyze LA insertion in the international arena becomes relevant, contributing to the visualization of the main producing countries and the dialogues established among them. Among the methodologies for assessing science and the dialogues established by authors, scientific collaboration analysis is highlighted.

Scientific collaboration is understood as the joint work of researchers, sharing data, equipment and ideas on a project aiming at producing scientific knowledge, and points co-authoring as an indicator of this activity (Katz & Martin, 1997). Scientific Collaboration reflects a range of exchanges among researchers and is a procedure that optimizes the production of science.

Scientific Collaboration among authors or institutions implies an alliance of assumptions and core objectives of a project, the establishment of a division of labor, the interaction between researchers, information sharing and coordination of these different relationships of joint investment (Olmeda Gómez, Perianez-Rodriguez & Ovalle-Perandones, 2008).

Scientific collaboration brings benefits to research, especially when it is held internationally, in order to increase productivity, visibility and impact, thereby being encouraged and supported by researchers and funding agencies (Glänzel & Schubert, 2005).

Collaboration frequencies show the connectivity between researchers, institutions or countries. However, they contain less information than the normalized co-authorship index, such as Salton's Cosine (SC), which consistently reveals issues related to intellectual structure of an area, when it relates the differences of scientific performance of researchers, institutions or countries, which varies among areas.

Salton's Cosine measures the proximity of two authors, regardless of the total number of scientific publications produced. Salton's Cosine (SC) is defined as the ratio between the frequency of co-authorship of two authors (X and Y) and the square root of multiplication of the number of articles produced by X and Y. The mathematical expression is given by:

$$SC = \frac{coauth(X, Y)}{\sqrt{auth(X) \cdot auth(Y)}}$$

where:

coauth(X, Y) = number of articles produced in co-authorship by authors X and Y;

auth(X) = total of papers produced by author X;

auth(Y) = total of papers produced by author Y.

This index standardizes co-authorship values in different areas of knowledge by presenting numbers ranging from zero to one: the higher and closer to one, the more intense the scientific collaboration between the authors; the closer to zero, the less intense the cooperation between them, in the light of the total scientific production of these authors. Zero for SC means the absence of co-authorship between the two authors; a value equal to one indicates that all the scientific production of the two authors was developed in co-authorship.

Luukkonen et al. (1993) stands out the importance of using both absolute and normalized indexes, considering that each of them brings a different kind of information for understanding the scientific proximity. Absolute indexes respond to questions concerning betweenness and proximity in networks; and normalized indexes represent the intensity of the relationship between the pairs.

This paper aims to conduct a comparative analysis of scientific collaboration proximity trends generated from absolute indicators and indicators of collaboration intensity in the field of Environmental Sciences in Latin America (LA), in order to identify possible existing biases in the absolute indicators of international cooperation, due to the magnitude of scientific production of these countries in mainstream science. In this way, we seek to contribute to the visualization of real collaboration intensities between countries, and for reflection on the challenges posed to S&T indicators in peripheral geographical regions, which may be inadequately described by absolute indicators. We adopted the notion of periphery as composed of elements having a lower status in a domain with unequal participation of its members.

More specifically, the objective is to analyze the compared forms of absolute and normalized values of co-authorship among Latin America countries and their main collaborators, in order to observe similarities and differences expressed by two indexes of frequency in relation to scientific collaboration trends in LA countries. In addition, we aim to visualize and analyze scientific collaboration networks with absolute and normalized indexes of co-authorship through SC among Latin America countries and their collaborators, comparing proximity evidenced by two generated collaborative networks - absolute and relative indicators.

The area of Environmental Sciences is justified as the research universe of analysis, given the social relevance these studies have gained currently, and the lack of studies in the literature to examine their scientific production trends.

METHODOLOGY

In this research, we used the theoretical and methodological contribution of bibliometric indicators of production and scientific collaboration, as support for visualization and analysis of intensity of scientific collaboration relations in LA, in the area of Environmental Sciences, indexed in Scopus.

Data collection, conducted in January 2016, comprised a period of 10 years (2006-2015) for the 18 AL countries that produce knowledge, using as a search term in the "Advanced search" option, the expression:

SUBJAREA(ENVI) AND AFFILCOUNTRY(Belize or "Costa Rica" or "El Salvador" or Guatemala or Honduras or Nicaragua or Panama or Argentina or Bolivia or Brasil or Brazil or Chile or Colombia or Ecuador or Paraguai or Peru or Uruguai or Venezuela or Mexico) AND PUBYEAR aft 2005 AND PUBYEAR bef 2016 and DOCTYPE(AR)

For all of the LA countries, we retrieved 45,439 articles in the area of Environmental Sciences in the 2006-2015 period. The analysis was restricted to countries: Brazil, Mexico, Argentina, Chile and Colombia as they produced 42,805 articles, corresponding to 94% of total production, a percentage considered representative and significant for this study.

For each of these five countries, we limited the search term to the name of each of them, in order to retrieve their key collaborating countries, with their respective co-authorship frequencies, in the analyzed period.

Then, we verified the co-authorship frequencies among the five countries and their key collaborators, common to all five countries, totaling 11 collaborating countries, namely: United States, United Kingdom, France, Spain, Germany, Canada, Netherlands, Australia, Italy, Belgium and Switzerland.

We built the 5x11 matrix (5 LA countries and 11 collaborators) with absolute frequencies of co-authorship among them. We verified the total number of articles published by each collaborator country in order to calculate the SC. We built the 5x11 matrix with the indexes of co-authorship normalized through SC.

Next, we generated the scatter plot from the 55 pairs of absolute and normalized frequency values, adopting as a crossing point of the axes X (absolute frequencies of co-authorship) and Y (normalized frequency through SC), the median value of respective indexes, as follows:

median of absolute frequencies equal to 177 articles and median of relative frequencies (SC) equal to 0.01342. We then analyzed the pairs of values present in the four generated quadrants. To support this analysis, Pearson correlations were calculated for joining quadrants on trends similarities. Thus, the correlations were calculated for the quadrants Q1 and Q3, jointly, and Q2 and Q4, also jointly.

Then, we generated two egocentric networks of scientific collaboration - absolute frequencies and normalized frequencies through SC using Pajek software. We proceeded to the analysis and comparison of similarity trends and differences between the two scientific collaboration indicators.

ANALYSIS OF RESULTS

Tables 1a and 1b show the absolute frequencies and normalized through SC among the 5 LA countries and 11 collaborating countries, with absolute frequencies ranging between 42 and 2218 articles, and normalized index ranging between 0.01 and 0.04.

Table 1a. Matrix of absolute frequencies of co-authored articles

	United States	United Kingdom	France	Spain	Germany	Canada	Netherlands	Australia	Italy	Belgium	Switzerland	Papers total
coauthorship papers												
Brazil	2218	817	723	650	642	449	345	342	283	150	148	22236
México	1883	325	284	743	214	302	149	130	126	71	64	9778
Argentina	791	191	167	470	257	185	63	140	158	39	78	3972
Chile	768	263	198	473	291	177	68	148	99	64	70	6344
Colombia	459	157	124	308	114	75	73	77	42	44	72	1833
Papers total	189484	56965	34290	35381	46577	42059	20445	36348	28458	10411	14357	

Table 1 b. Matrix of normalized frequencies through SC of co-authored articles

	United States	United Kingdom	France	Spain	Germany	Canada	Netherlands	Australia	Italy	Belgium	Switzerland
coauthorship papers											
Brazil	0,034	0,023	0,026	0,023	0,020	0,015	0,016	0,012	0,011	0,010	0,008
México	0,044	0,014	0,016	0,040	0,010	0,015	0,011	0,007	0,008	0,007	0,005
Argentina	0,029	0,013	0,014	0,040	0,019	0,014	0,007	0,012	0,015	0,006	0,010
Chile	0,022	0,014	0,013	0,032	0,017	0,011	0,006	0,010	0,007	0,008	0,007
Colombia	0,025	0,015	0,016	0,038	0,012	0,009	0,012	0,009	0,006	0,010	0,014

Figure 1 shows the scatter plot of the observed values in Matrices 1a and 1b, with four quadrants defined according to the medians of each of the matrices.

In the first quadrant (Q1), with 22 pairs of co-authorship indexes, we observe values above the median for absolute frequency and for normalized indexes of co-authorship, corresponding to those collaborations that stand out both in quantity and intensity, i.e., visible by any absolute or normalized value analysis.

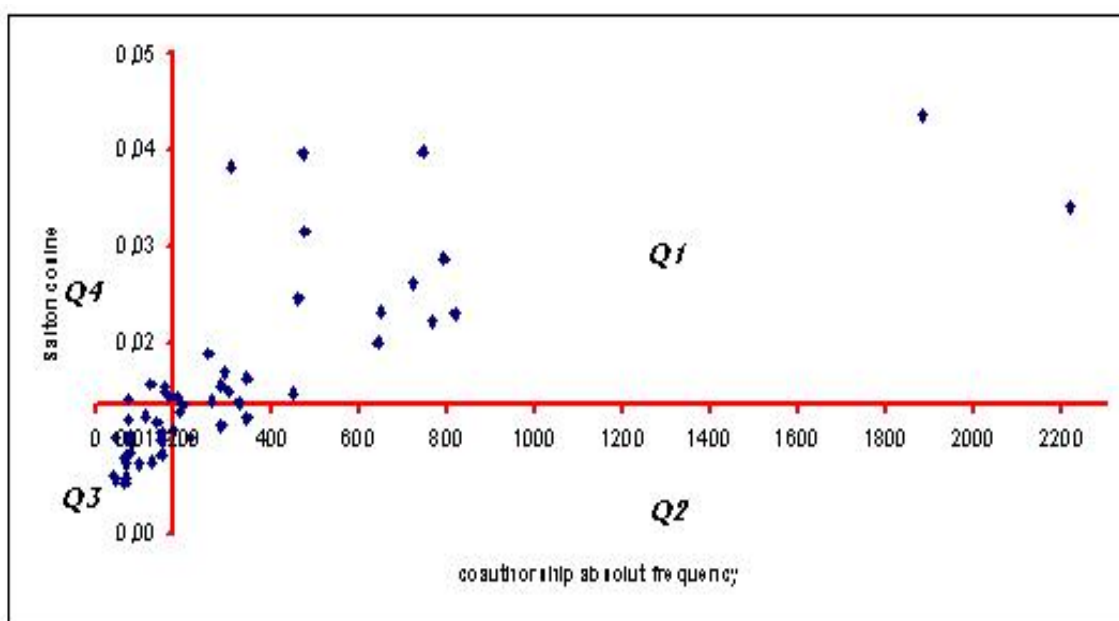
In Q2, with 5 pairs of co-authorship indexes corresponding to values above median for absolute frequency, and below median for normalized indexes of co-authorship, corresponding to collaborations with high frequency, however, poorly significant when observed in relation to the total of the scientific production of the countries involved.

In Q3, with 23 pairs of co-authorship indexes corresponding to values below median for absolute frequency and normalized indexes, relating to collaborations without emphasis in absolute and normalized form, that is, not perceptible by both analyzes.

In Q4, with 5 pairs of co-authorship indexes, with values below median for absolute frequency and above median for normalized indexes, there are 12 pairs of values, and 6 of them associated to Colombia and 4 to Argentina, corresponding to a low frequency of co-authorship of these countries with their collaborators, however, corresponding to an intense scientific collaboration when contextualized in the light of the total production of these countries. This collaboration intensity is not visible in an analysis that holds only to absolute values of co-authorship.

Thus, Q1 and Q3 totaled 45 pairs of indexes, corresponding to 82% of total pairs, and show the same collaboration trends both in absolute values and normalized through SC. The linear correlation coefficient was calculated among these 45 pairs of indexes, which resulted in $r = 0.76$, indicating a strong positive trend of correlation between the two indexes (absolute and normalized).

Figure 1: Scatter plot between absolute e normalized frequency



On the other hand, Q2 and Q4 totaled 10 pairs of collaboration index and point out distinct and reverse trends of collaboration significance among the analyzed LA countries with key collaborating countries. This behavior may be due to the fact that scientific cooperation be more significant for the scientific development of LA countries than the observed absolute frequency (Q4). Moreover, because the intensity is relatively low (in terms of SC) due to high production of scientific countries in the area, although the absolute frequency of collaboration is high. The linear correlation coefficient was calculated among the 10 pairs of indexes of these two quadrants, resulting in $r = -0.66$, indicating a relatively strong negative correlation between the two indexes.

Figures 2a and 2b show the scientific collaboration networks among the five LA countries and the 11 collaborating countries, generated from the matrices present in Tables 1a and 1b, in their absolute frequencies values (Figure 2a) and normalized through SC (Figure 2b). The red circles correspond to the LA countries and the blue ones to 11 key collaborators common to them all. The areas of the circles are proportional to the total scientific production of the countries.

Analyzing the network generated from the absolute values matrix, the higher frequencies of co-authorship occur between United States and 4 of the 5 LA countries, namely: Brazil, Mexico, Argentina and Chile, where the connecting threads between the nodes are thicker, especially between United States and Brazil. In addition, with higher frequencies of co-authorship, Brazil and United Kingdom, Brazil and France, Mexico and Spain.

Figure 2a: Co-authorship network, per absolute frequencies, among the LA analyzed countries

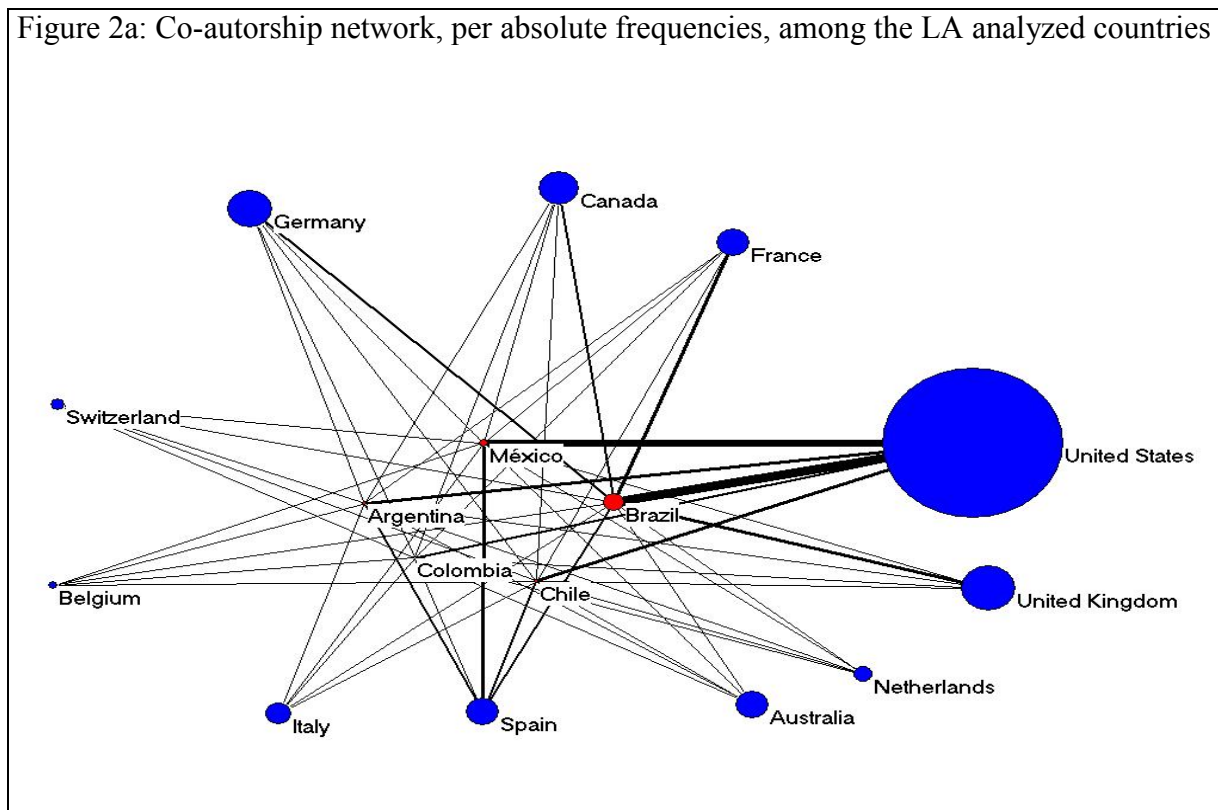
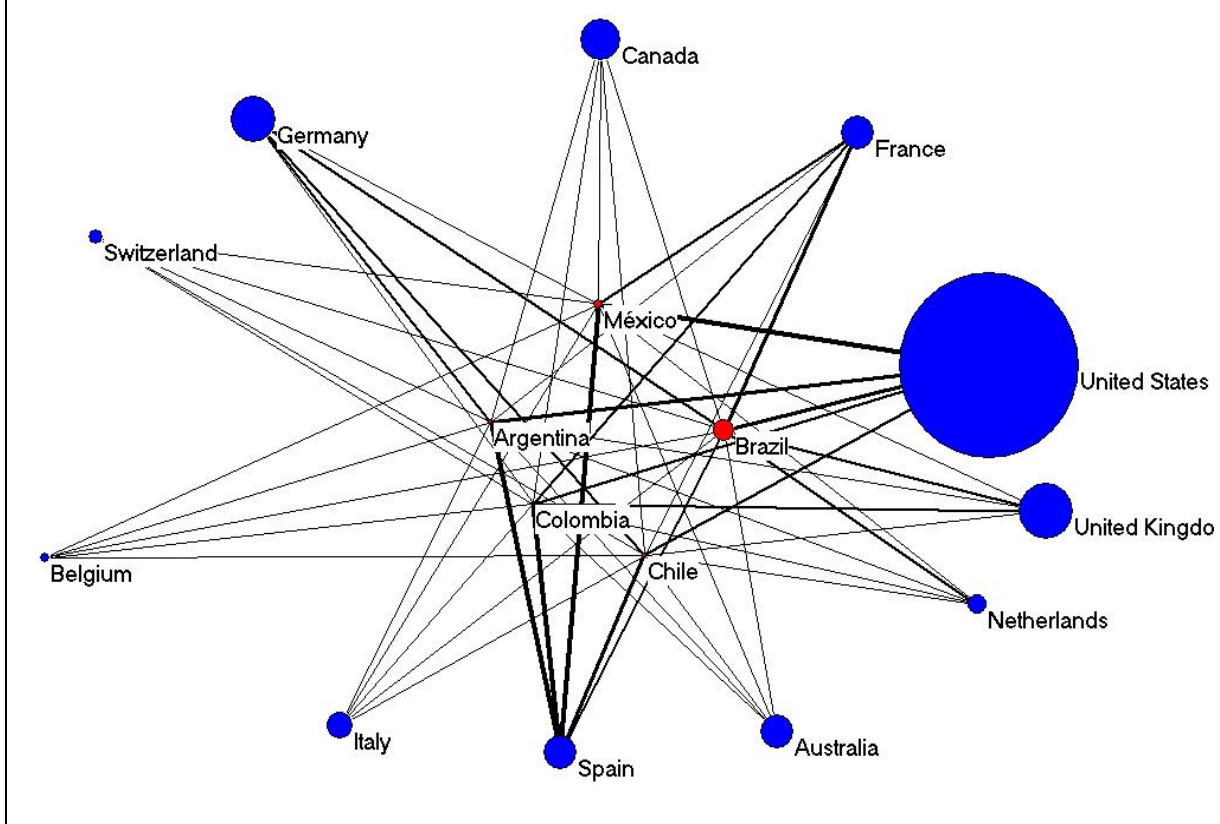


Figure 2b: Co-authorship network, per normalized frequencies through SC, among the LA analyzed countries



Regarding the generated network from the normalized values, the stronger co-authorship connections are highlighted between Spain and four LA countries, namely: Mexico, Argentina, Chile, Colombia and Brazil with less intensity, indicating that Spain has strong partnerships with LA countries, cooperation that are not visible on the co-authorship absolute frequency network. It has been hypothesized that this behavior is due to the linguistic proximity between them, considered a facilitating element of scientific cooperation. Brazil, Mexico and Argentina, with strong connections to the United States, confirming the cooperation trend among these countries, as observed in absolute values of co-authorship frequency. Also highlighting normalized co-authorships between Brazil and France and between Mexico and France, this latter not visible in the absolute frequencies network. In addition, other non-visible collaborations in Figure 2a between Germany and Chile, and between Germany and Argentina.

We also observe some discrepancies between networks of absolute and normalized values. Although in absolute terms, the co-authorships between Brazil and Spain are more than twice the frequency of co-authorship between Spain and Colombia, as shown in Table 1a, when we relativize them according to the volume of scientific production of the involved countries, intensities of scientific cooperation relationships, are reversed, i.e., the normalized intensity between Colombia and Spain (greater SC index of scientific collaboration, equal to 0.04) is twice the index achieved between Brazil and Spain (0.02). The absolute relationship between Colombia and the United Kingdom is among the less intense ones (157) and when we relativize this frequency in relation to scientific production of these countries, the collaboration index appears more significant (0.015).

CONCLUSIONS

Based on the results, the analysis of scientific collaboration through normalized indexes is considered essential for countries of great scientific production, since there are co-authorships that are not significant for the development of scientific production these countries, when examined in the light of their total production. Such as collaborations between Mexico and Germany, Brazil and Italy, Brazil and Australia, Argentina and the United Kingdom, and Chile and France.

Still, these normalized indexes show is significant in particular for understanding the scientific development of countries with lower production, considered in this study as peripheral, as the normalized indexes make these countries' intense collaborations visible, not identifiable in absolute terms, given the magnitude of other collaborations from countries with high production, even if they are not significant in the context of the latter (countries with large production in the area), for example, the collaboration between Colombia and Switzerland, Argentina and France, Argentina and Italy, Colombia and UK, Colombia and France.

From the results, we observed the need for absolute and normalized indicators to describe the scientific collaboration phenomenon in a more thoroughly way, once these indicators provide complementary information.

REFERENCES

- Gazzola, A. L.& Didriksson, A. (2008). *Tendencias de la educación superior en América Latina y el Caribe*. Unesco. Iesalc.
- Glänzel, W.& Schubert, A. (2005). Analysing scientific networks through coauthorship. In H.F. Moed , W. Glänzel, &U. Schmoch (Eds). *Handbook of quantitative science and technology research*. Netherlands: Springer.
- Gutierrez, R.C.C.& Gracio, M.C.C. (2015). Indicadores de produção científica na América Latina: um estudo na área da Matemática. In: J. Zabala Vázquez; R. Sánchez Jiménez & M. A.García Moreno. (Ed.). *Desafíos y oportunidades de las Ciencias de la Información y la Documentación en la era digital: actas del VII Encuentro Ibérico EDICIC 2015*. 7ed. Madrid: Universidad Complutense de Madrid.
- Katz, J.S. & Martin, B.R.(1997). What is research collaboration? *Research Policy*, 26, 1-18.
- Luukkonen, T. et al.(1993). The measurement of international scientific collaboration. *Scientometrics*, 28, 15-36.
- Olmeda Gómez, C., Perianez- Rodriguez, A.& Ovalle-Perandones, M. A. (2008). Estructura de las redes de colaboración científica entre las universidades españolas. *Ibersid 2008: revista de sistemas de información e comunicación*, p.129-140.
- Velho, L. (2005). S&T Institutions in Latin America and the Caribbean: an overview. *Science and Public Policy*, 32(2), 95-108.

POSTER SESSION 1

**Indicators, Assessment,
Funding and Innovation**



The research activity index at the Universitat Politècnica de València (IAIP): How an institution can complement national regulation on the productivity of university professors in research and teaching activities

J.A. Conejero, J.E. Capilla, L.M. Sánchez-Ruiz, V. Amigó, A. Blasco, V. Botti, J.C. Cano, J. Capmany, A. Chiralt, S. Cuesta, J.M. García Álvarez-Coque, C. Jordá, J. Juliá, P. Miguel, M.A. Miranda, and A. Peris

aconejero@upv.es, jcapilla@upvnet.upv.es, lmsr@mat.upv.es, vamigo@mcm.upv.es, ablasco@dca.upv.es, vbotti@dsic.upv.es, jucano@disca.upv.es, jcapmany@iteam.upv.es, dchiralt@tal.upv.es, sacuesta@esc.upv.es, jmgarcia@upvnet.upv.es, cjorda@upv.es, jfjulia@esp.upv.es, pmiguel@cst.upv.es, mmiranda@qim.upv.es, aperis@mat.upv.es

IUMPA – IIAMA – CITG – ITM – ICTA – DSIC (GTI-IA) – DISCA (GRC) – iTEAM – IIAD – DE (L. Luz) – DECS (GEID) – DCMA – CEGEA – ICITECH – ITQ – IUMPA.
Universitat Politècnica de València, Camino de Vera s/n, Valencia, E-46022 (Spain)

ABSTRACT

The Universitat Politècnica de València (UPV) created an index of research activity in 1991, with the goal of being an objective criterion to distribute internal University funds for fostering the research activity of its professors. In 1998-2000 the University approved the current index regulation named Personalized Research Activity Index (IAIP). This index has arrived until the present day, with minor modifications.

Since 2012, it has also been used to complement the national regulation on the teaching activities of the university professors at the UPV. Reductions on the teaching workload of the professors can be added to the ones stated in the national regulation attending to that inner evaluation of the research activity.

In this communication reports the work in progress of the group that is reviewing the current index assessment. We will show here the master lines of this revision with respect to the research production. Future lines of work are also exposed.

INTRODUCTION

In 2012, the Spanish Government approved some measures in order to rationalize the budget given to education (BOE, 2012). In particular, it was approved a redefinition of university professors workload in terms of the positive evaluation of six-years research periods (*tramos de investigación* or *sexenios*, in Spanish).

The evaluation of these research periods is mainly based on research papers published in journals indexed in Web of Science. In that case, the evaluation takes into account the position of the journal the year in which each paper is published. In order to attend to the singularities of each field, additional criteria are stated in order to consider other research items such as research books, book chapters, contributions published in conference proceedings of relevant conferences in the area, or artistic production. The last version of these criteria was published in (MECD, 2015). For being evaluated, each lecturer, teacher or professor proposes his/her 5 most significant research achievements during a six-year period. The evaluation of the period can only be either positive or negative, with no distinctions of achievement levels when it is positive.

Attending to the aforementioned legislation, the government of each university can only take one of the following two decisions for determining the teaching workload of its professors: The first option is to set all the professors workload equal to 24 ECTS¹. The second one is to set differences according to the existence of six-years research periods recently achieved: 32 ECTS for professors that do not conduct research, 24 ECTS for the ones with a six-year evaluation achieved in the last six years, or 16 ECTS for professors that have additionally achieved several consecutive periods of positive evaluations (3 periods for associate professors and 4 periods for full-professors).

The Universitat Politècnica de València created a research activity index in 1991, with the goal of being an objective criterion for distributing University funds gear to foster the research activity of its professors. In 1998-2000 the University approved the current index regulation, named Personalized Research Activity Index (IAIP)². With minor modifications, this index has been used for the last 15 years, but in the last years a revision of it has been seen as necessary.

The IAIP was conceived for fostering the research and technology transfer activity, and to promote the efforts for capturing external financial resources. It has been used in order to assign internal budgets to professors and researchers for first and interdisciplinary research projects, for stimulating the participation in conferences, for staying abroad during some periods of time, and for awarding research grants to PhD students. Additionally, the annual budget of departments, research institutes and other smaller research centres depends on the accumulated IAIP of the researchers integrated in each structure.

But since 2012, it has also been used to complement the aforementioned national regulation on the teaching workload of university professors at the UPV. Depending on the evaluation obtained in that index by a professor, additional reductions of his/her teaching workload can be added to the one stated by the national regulation. Ultimately, this allows the institution to redirect the activity of their professors according to its goals and singularities.

As we will see later, the use of such an index presents some limitations in order to conduct a correct evaluation of the research activity of every single university professor. However, attending to the singularities of the Spanish university system and its governance, the use of such an index presents some advantages for the management and distribution of resources within the members of the institution.

This communication reports the work in progress of the group that is reviewing the current assessment index IAIP. We will show here the master lines of this revision with respect to the research production. It is pending to consider the fund raising for research activities. Further information can be obtained from the website of the Vice-Rectorate for Research, Innovation and Transfer of the UPV.

¹ We point out that the misuse of the concept of ECTS is taken directly from the law.

² Índice de Actividad Investigadora Personalizado.

THE IAIP INDEX OF THE UPV

This IAIP¹ is a quantitative indicator set to measure the amount of research activity of every single university professor of the UPV. Every year professors submit their contributions which are later validated in order to be computed by the indicator. To compute the value of the IAIP indicator, points are assigned to every single contribution of a professor. Each professor receives the sum of points given to his/her contributions. In order to compute the indicator of a research structure (a research institute or a department), one has to add up the contribution of all their members. An aggregated and weighted index that accumulates the activity during the last 4 years is usually considered (VAIP)³. It is much more practical to use this last index since it presents smaller variations respect to the IAIP.

In general, every professor receives the following amount of points for each research item, as long as the number of co-authors of the contribution is smaller or equal to 4. Only if there are 5 or more, the number of points received by every single co-author of a contribution is reduced proportionally to the number of authors beyond 4.

Research papers in journals

University rankings such as ARWU (ARWU, 2016) or Leiden (CWTS – Leiden Ranking, 2016) recognise the production of research papers that are listed in journals appearing in the Web of Science database produced by Thomson Reuters. A special mention is given to papers appearing in the top 20% (ARWU – Rankings by area) or top 1,10 or 50% (Leiden). For every publication we consider the highest position of the journal in which it appears among all the possible categories of the Journal Citation Reports. Papers appeared in journals among the 10% most cited are to received 15 points. If they are in the rest of the first quartile, 12 points. Papers in the second quartile 9 points, in the third one 9 points, and in the fourth, 4 points. A publication in a journal not indexed by Web of Science receives only 1 point. A special bonus of 100 points is given to the first author of papers published in Science or Nature, and 50 points for the rest of the authors.

We point that for areas of Economy, Humanities, and Architecture, other databases are considered instead of Web of Science, such as Scopus (with the SCI-mago JCR), ERIH, In-RECH, MIAR, and the Avery Index.

Additionally, if a professor is member of the editorial committee of a journal, then he/she receives every year as many points as the points corresponding for a publication on that journal. In case, that he/she was the editor in chief, then he/she receives as many points as two publications there.

Research papers in proceedings of conferences

Authors of research papers published in proceedings of international conferences receive 1 point per item, and 0.5 for the case of national conferences. A special mention is done to conferences listed in the CORE conference ranking of conferences in ICT: 8 points are assigned to authors of a publication in a conference in the A+ list, 4 in the case of the A list,

¹IAIP stands for Índice de Actividad Investigadora Personalizado

³VAIP stands for Valoración de la Actividad Investigadora Personalizada. The VAIP is computed as the sum of the IAIP of the last year, plus 0.75 times the evaluation of the year before, plus 0.5 times the one of two years before, plus 0.25 times the one of three years before.

and 2 when the conference is classified as B. The values assigned for the membership to the program or scientific committee of a conference, as long as there is also a participation as author or co-author of a communication, are 2 points in case of international conferences and 1 point for national conferences. A maximum of 10 points can be obtained in this dimension. Apart from this, in case that the membership is as general chair or president of the scientific committee of a conference, then the points received are multiplied by two, with a maximum of 10 points in this case, but never getting beyond 15 points for committee memberships in total.

Books and monographs

The publication of a book is considered equal to the publication of a research paper, with an amount of points ranging from 12 to 1 depending on the importance of the publishing house. If the publication is not of a whole book, but of chapters, the author receives half of the points. The points received by the authorship of chapters cannot exceed what could be obtained for a whole book. As a reference, the Scholarly Published Indicators are partially taken into account for classifying the editorials, and then for evaluating the books attending to the relevance of the editorial that publishes it.

Artistic production

For the field of Fine Arts, points for the participation in an exposition are awarded, too. An expert committee of professors of the University approves a list of expository places in order to conduct the evaluation of this participation. Such a list is updated at least once every four years. The exhibitors are classified as A,B or C depending on the importance considered by that committee. The points received by a professor for the participation on a exposition depend on the category assigned to the exhibitor. If classified as A, then 9 points are assigned per contribution, and 6 and 4 in case of spaces labelled as B or C, respectively.

Prizes and recognitions

If the prize is perceived as recognition of a career, then the points are assigned every single year. There is a committee that assigns between 5 to 25 points to each received prize depending on an evaluation of the prize itself and related to other prizes already considered for evaluation.

If the contest is framed into Architecture or Design, which is quite frequent, the number of points assigned can be compared with a research publication. Again, an inner committee evaluates the relevance of the prize and assigns some points to the professor for the achievement.

CONCLUSIONS AND FUTURE LINES OF WORK

As we have seen the IAIP is a quantitative indicator to obtain a weighted measured of the amount and quality of the R&D and technology transfer activities. However, there are several improvements that are going to be incorporated in the index calculation. Progress have been done towards the incorporation of the following quality criteria, although no particular measures have been taken yet:

- Better consideration of publications of very high impact in order to stimulate high quality papers in the best journals.
- More precise evaluation of contributions in conferences and congresses in those research areas where these contributions have higher impact.
- More recognition of research books, currently poorly considered due to the lack of objective criteria.
- The degree of internationalization, for fostering the collaboration with top institutions.
- The interdisciplinarity of the research and the collaboration of different areas.
- Incorporating specific criteria applicable to specific research areas (e.g. sciences, philology, fine arts, etc. require different quality criteria).

By now, the impact of a publication is understood to be given by the impact factor of the journal in which it appears. The consideration of citations, of other metrics, as a measure of the impact of a single contribution or as a measure of the relevance of the academic trajectory of a professor should be also taken into account in a short future.

There is also the agreement in stimulating the participation in funded projects, and promoting all tech transfer activities. By now, it was given 1 point to a professor for the mere participation in a research project, and 2 additional points if the participation was as principal investigator or work package leader of the project. Additionally, a number of points are assigned to the researchers participating in the project according to the budget. These elements and their quantification are still under revision.

We also want to mention that other activities to be fostered are the creation of spin-offs and the international patent registration. In both cases, it is expected that the number of points assigned was at least the points given to a publication in a journal in the top 10%.

The implementation of an index such as the IAIP permits to give a one-dimensional result of the evaluation of the quantity and quality of the research conducted by every single professor at the institution, despite it is not as significant and rich as a peer evaluation of the research activity of a professor by experts of the same research field. However, our system permits to consider many more aspects than what is taken into account in the 6-years evaluation periods (*sexenios*), and it also permits to better recognise the research activity in areas where the computation of the impact factor of a publication present some limitations.

Finally, we also want to indicate that such a system can be implemented in any university. The weights given to each dimension can be change in order to align them with the research policy fostered by the management team of each institution.

REFERENCES

ARWU – Academic Ranking of World Universities (2016) Methodology.

<http://www.shanghairanking.com/ARWU-Methodology-2015.html> Last access on March, 29th, 2016.

CWTS – Leiden Ranking (2016) Methodology.

<http://www.leidenranking.com/methodology/datacollection> Last access on March, 29th, 2016.

Jefatura del Estado. *Real Decreto-ley 14/2012, de 20 de abril, de medidas urgentes de racionalización del gasto público en el ámbito educativo*. Boletín Oficial del Estado. April 21st, 2012 (pp. 30977-30984).

Ministerio de Educación Cultura y Deporte - MECD (2015). *Resolución de 26 de noviembre de 2015, de la Comisión Nacional Evaluadora de la Actividad Investigadora, por la que se publican los criterios específicos aprobados para cada uno de los campos de evaluación*. Boletín Oficial del Estado. November 30th, 2015 (pp. 113060-113078).

UPV - Vice-Rectorate for Research, Innovation and Transfer of the UPV (2016) Research activity index. <http://www.upv.es/entidades/VIIT/info/216020normali.html> Last access on March, 29th.



Bibliometric indicators and activity scores for digital scholars

Susanne Mikki* and Marta Zygmuntowska**

* susanne.mikki@uib.no

University of Bergen Library, Norway

** marta.zygmuntowska@uib.no

University of Bergen Library, Norway

Abstract

The use of academic profiling sites is becoming more common, and emerging technologies boost researchers' visibility and exchange of ideas. In our study we compared profiles at five different profiling sites. These five sites are ResearchGate, Academia, Google Scholar Citations, ResearcherID and ORCID. The data set is enriched by demographic information including age, gender, position and affiliation, which are provided by the national CRIS-system in Norway (CRISTin). We investigated the correlation between bibliometric measures, such as publications and citations, and user activities, such as downloads and followers. We find different bibliometric indicators to correlate strongly within individual platforms and across platforms. However, there is less agreement between the traditional bibliometric and social activity indicators.

Findings

Usage of academic network sites and demographic distribution

We find that approximately 37% of researchers at the University of Bergen have at least one profile, the prevalence being highest (> 40%) for members at the Faculty of Psychology and the Faculty of Social Sciences. Across all disciplines, ResearchGate is the most widely used platform. Researchers are reluctant to maintain multiple profiles, and there is little overlap between different services.

Our study confirms established hierarchical patterns in regard to age, gender and position, showing the elder, male professors are best represented. However, this result is not unambiguous. We find also that post doc fellows to a large extent embrace these sites, which reflects their stronger need to be visible on the job market.

Correlation of traditional metrics between network sites

We used Spearman's rank correlation and compared the traditional bibliometric indicators such as number of publications, citations and h-index and find these indicators strongly correlated within and across platforms.

Correlation of traditional metrics and altmetrics

For many years researchers have been evaluated by their publication productivity and citation impact. However, academic profiling services may provide a new way of measuring scholarly impact. We divided the provided alternative metrics into two groups: Publication score and

Social activity. The first group is related to publications, and their impact measured by e.g. downloads, views. The second is related to the users' social activity and online network engagement.

We find less agreement between traditional bibliometric indicators and indicators for social activity. We find some agreement between the traditional metrics and the publication scores.

Our results are consistent with previous findings showing that social network sites reflect the same hierarchical structures as in real life (Menendez, de Angeli, & Menestrina, 2012). Scientists that publish a lot, and probably have a longer career and higher position, get more followers but do not necessarily follow others. Although network sites gain ground, their uptake is far from universal, and available metrics should therefore be used carefully in an evaluation context. In addition, data manipulation is an issue to be aware of and looked into.

Menendez, M., de Angeli, A., & Menestrina, Z. (2012). Exploring the virtual space of academia *From Research to Practice in the Design of Cooperative Systems: Results and Open Challenges* (pp. 49-63): Springer.

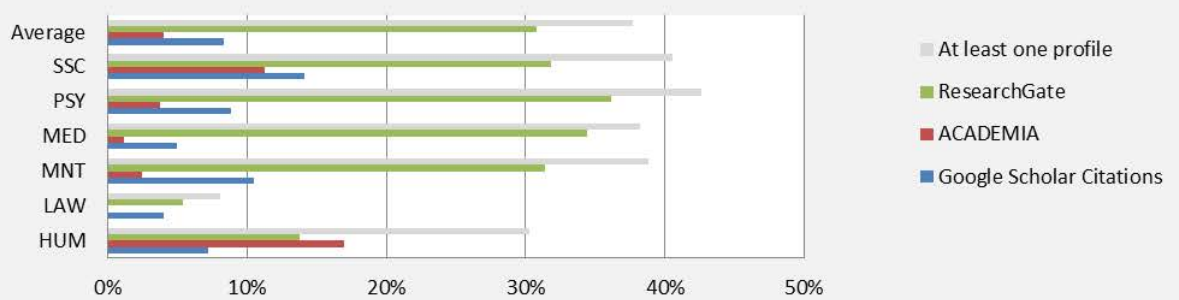
Impact and activity scores of digital scholars

Susanne Mikki and Marta Zygmuntowska

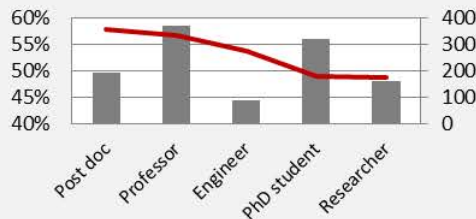
University of Bergen Library, Norway

Presence

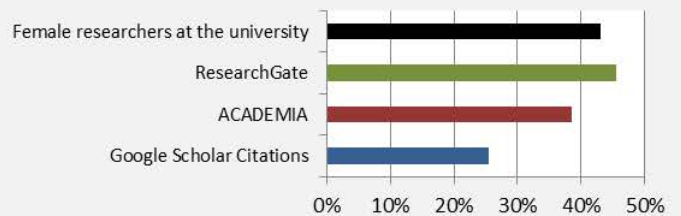
We find an over average presence of scholars at the faculties of Psychology (PSY) and Social Sciences (SSC).



We find the highest presence among post docs and professors

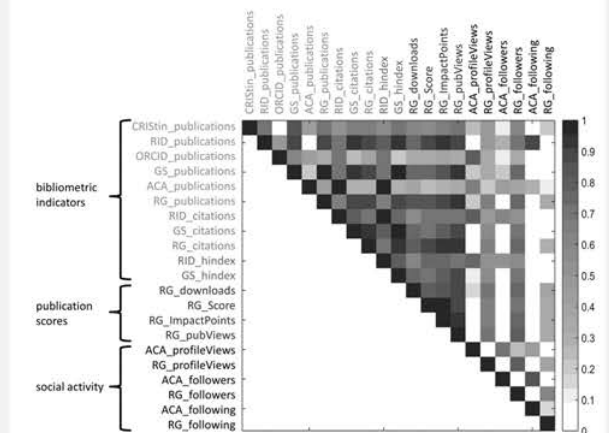


We find the female presence lower than the male, highest at ResearchGate and lowest at Google Scholar Citations



Indicators

We find the traditional bibliometric indicators to correlate **strongly** within and across sites. However, there is *less* agreement between the traditional bibliometric indicators and the indicators for social activity.



Mikki, S., Zygmuntowska, M., Gjesdal, Ø. L., & Al Ruwehy, H. A. (2015). Digital Presence of Norwegian Scholars on Academic Network Sites—Where and Who Are They? *Plos One*, 10(11), e0142709. doi: 10.1371/journal.pone.0142709



Mapping scientific controversy in Twitter: the Maya city hoax¹

Elena Denia *

*elenadenia@ingenio.upv.es
INGENIO (CSIC - UPV), Ciudad Politécnica de la Innovación
Camino de Vera, s/n 46022 Valencia (Spain)

ABSTRACT

The present poster reflects a study on the social diffusion of science and public attitudes toward science taking advantage of the available data of the online social network Twitter (real-time nature) and focused on a specific science new that turned out to be a hoax. I consider two lines of study of different nature: (1) on the one hand the aim is to offer insights into to what extent the structure of the network influences the information spread and serves to capture public attention, as well as identify common features of the major influencers; (2) on the other hand I carry out a deeper analysis concerning to the content of the message – tweet–, by using data mining technics, with the purpose of exploring the main elements that play a key role in terms of laypeople interest, trust and engagement, and to observe the predisposition to change opinion while dealing with a fallacious case in which there was no scientific evidence.

The particular case of study is the Maya city controversy. In May 2016, a 15-year-old schoolboy William Gadoury, from Quebec, Canada, compared maps of 22 star constellations to the ancient Maya with Google Earth images of Mexico's Yucatan peninsula. Despite there was no scientific evidence, general media spread a new saying that William found a lost Maya city. We study the reception of such a false new.

Our hypothesis is that the image of science and public opinion of scientific facts depend both on the network structure and on the content of the information. What information can we get from these two approaches of such a different nature?

INTRODUCTION

Both academia and the public authorities advocated that greater permeability between science and society today is a desirable and even essential objective. With the rise of digital social networks, that has led to virtual communities sustained in an architecture of participation, it seems reasonable to investigate novel or insufficiently studied aspects of the relationship between science and society and the public image of science. That is to consider innovative tools for measuring the social perception of science, a social aspect studied extensively over the years by *Eurobarometer* surveys in Europe, by *Fecyt* reports in Spain, among others.

The present poster reflects a study on the social diffusion of science and public attitudes toward science taking advantage of the available data of the online social network Twitter (real-time nature) and focused on a specific science new that turned out to be a hoax. I consider two lines of study of different nature: (1) on the one hand the aim is to offer insights into to what extent the structure of the network influences the information spread and serves to capture public attention, as well as identify common features of the major influencers; (2)

¹ This work was supported by Ministry of Economy and Competitiveness (Grants for predoctoral contracts for the training of PhD candidates 2014 and Proyecto EXTRA, CSO2013-48053-R).

on the other hand I propose a deeper analysis concerning to the content of the message – *tweet*–, by using data mining technics, with the purpose of exploring the main elements that play a key role in terms of laypeople interest, trust and engagement, and to observe the predisposition to change opinion while dealing with a fallacious case in which there was no scientific evidence.

The particular case of study is the Maya city controversy. In May 2016, a 15-year-old schoolboy William Gadoury, from Quebec, Canada, compared maps of 22 star constellations to the ancient Maya with Google Earth images of Mexico’s Yucatan peninsula. Despite there was no scientific evidence, general media spread a new saying that William found a lost Maya city. We study the reception of such a false new.

Our hypothesis is that the image of science and public opinion of scientific facts depend both on the network structure and on the content of the information.

METHODS AND TECHNICS

The analysis software used is R –and its packages–, a powerful computer tool used widely by data miners and statisticians, which is freely available under the GNU General Public License. Also, in order to have access to the data, it is necessary to use the Twitter API (Application Programming Interface).

In our particular case, to extract the data we took the keywords ‘maya city’, and saved it in a file with metadata of the tweets from 2016-05-10 to 2016-05-19, resulting in a sample size of 11,495 tweets. Again, for the next time period of 9 days, to see the decreasing interest on the issue over time, we collected 1,708 tweets. The metadata includes varied information of each tweet and also about users who tweeted.

To analyze the content of the tweets we use techniques of text mining, including text cleaning, topic modelling, sentiment analysis and word associations. On the other hand, to study the structure of the network and features of the agents involved, we use social network analysis (SNA) to build the network, identify major influencers and track the message propagation.

RESULTS

We must be careful and consider different things for this kind of analysis. For example, something surprising from the data scraping is that in the Russian tweets (or linked urls) the word 'мая' occurs, meaning the month May, pronounced as 'maja' (the я is 'ja' in Russian). This is picked up by the Twitter search algorithm for our keywords. We also collected Japanese tweets, but in this case they actually refers to the case of study. If we look at one of the major influencers ‘@newton_science’, it is the account of a science magazine in Japan.

Other major influencer, those who have received more retweets, replies to tweet, or have been mentioned in conversations, etc., in the Maya city controversy is ‘@darrenaronovski’, who is a film maker. This gives us information about agents and its social influence, further than general media.

The wordcloud reveals that users pay special attention on concepts like ‘teen’, ‘experts’, ‘discovers’ and ‘marijuana’. The last one refers to a joke, suggesting that such an image was not revealing a Maya city but a marijuana field.

CONCLUSIONS

- With the rise of digital social networks that has led to virtual communities it seems reasonable to investigate novel or insufficiently studied aspects of the relationship between science and society and the public image of science.
- The real-time nature of the study provides unexplored dimension of public opinion about science issues.
- To investigate the content of the tweet and the structure of the network we need two approaches of different nature, therefore our results are complementary but not comparable.

Figure 1: Screenshot: running the code in R.

```

48 ## TOP FREQUENT TERMS (funciona bien con el inglés, debido al paso del text
49
50 # inspect frequent words
51 (freq.terms <- findFreqTerms(tdm, lowfreq = 20))
52
53 term.freq <- rowSums(as.matrix(tdm))
54 term.freq <- subset(term.freq, term.freq >= 20)
55 df <- data.frame(term = names(term.freq), freq = term.freq)
56
57 library(ggplot2)
58 ggplot(df, aes(x=term, y=freq)) + geom_bar(stat="identity") +
59   xlab("Terms") + ylab("Count") + coord_flip() +
60   theme(axis.text.element_text(size=7))
61
62
6335 (Top Level) >

```

Environment History

Global Environment -

Data

- ret_clean2016.06_ 105 obs. of 1 variable
- ret.df 105 obs. of 16 variables
- ret22016.06.28 105 obs. of 16 variables

Values

apt "47571889-TJED0i3fD3ZQjArym7wx0kyAgaALk1Ck1BZ_

Files Plots Packages Help Viewer

R: Data Frames - Find in Topic

data.frame (base) R Documentation

Data Frames

Description

This function creates data frames, tightly coupled collections of variables which share many of the properties of matrices and of lists, used as the fundamental data structure by most of R's modeling software.

Usage

```

data.frame(..., row.names = NULL, check.rows = FALSE,
            check.names = TRUE,
            stringsAsFactors = default.stringsAsFactors())

```

Console ~/Desktop/Utrecht/RDataMining/

R es un proyecto colaborativo con muchos contribuyentes. Escriba 'contributors()' para obtener más información y 'citation()' para saber cómo citar R o paquetes de R en publicaciones.

Escriba 'demo()' para demostraciones, 'help()' para el sistema on-line de ayuda, o 'help.start()' para abrir el sistema de ayuda HTML con su navegador. Escriba 'q()' para salir de R.

[Workspace loaded from ~/Desktop/Utrecht/RDataMining/.RData]

> |

REFERENCES

Bauer, M. W., Allum, N., & Miller, S. (2007). What can we learn from 25 years of PUS survey research? Liberating and expanding the agenda. *Public understanding of science*, 16(1), 79-95.

Ceron, A., Curini, L., Iacus, S. M., & Porro, G. (2014). Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. *New Media & Society*, 16(2), 340-358.

Avellaneda, R. P. (2001). La cultura científico-tecnológica de las sociedades de la modernidad tardía. *Treballs de la Societat Catalana de Biologia*, 35-63.

Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E. P., Yan, H., & Li, X. (2011, April). Comparing twitter and traditional media using topic models. In *European Conference on Information Retrieval* (pp. 338-349). Springer Berlin Heidelberg.



Visibility and Impact of Research Data Sets in the Life Sciences supported by a Novel Software Infrastructure

Claudia Kramer^{*}, Nicole Jung^{**} and Pierre Tremouilhac^{***}

^{*} claudia.kramer@kit.edu

KIT Library, Karlsruhe Institute of Technology, H.-v.-Helmholtz-Platz 1, Eggenstein Leopoldshafen, 76344 (Germany)

^{**} nicole.jung@kit.edu

Institute of Organic Chemistry, Fritz Haber Weg 6, 76344 Karlsruhe (Germany)

^{***} pierre.tremouilhac@kit.edu

Institute of Toxicology and Genetics, Karlsruhe Institute of Technology, H.-v.-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen (Germany)

ABSTRACT

Research data management and publication processes in the life sciences are often underdeveloped in terms of automation and computation. We implemented an Open Source infrastructure which supports the experimentation, documentation and publication of research data in chemical laboratory research environments. Datasets can be published and are automatically citable via DOI. We examined how the workflow for the citation of data sets can be improved to increase their visibility and impact.

INTRODUCTION

The visibility of experimental data and their accessibility are critical factors for the single researcher and also project teams as those data form the basics for a progress of their work. In addition, the retrieval of research data is highly important for the wider scientific community to extend its shared knowledge. In this context, the efficient acquisition and the management of data are processes of high importance within the publication life cycle. Despite their relevance, the data management and publication processes are often underdeveloped in terms of automation and computation, which is especially true in the field of life sciences (Frey, Bird 2013). Incentives, such as the providing of equipment and workflows for the acquisition of research data, their storage and publication as well as citation of data can help to foster the culture of data sharing in the context of Open Science. (Nature Biotechnology 2009).

METHODS

We implemented a basic management software which supports the experimentation, documentation and publication of research data in chemical laboratory research environments. The new research infrastructure includes two main software modules designed for the management of research data: (1) The ELN (Electronic Laboratory Notebook) enables the planning of experiments and the documentation of research data, including a basic LIMS (Laboratory Information Management System) for data acquisition and handling. The ELN allows seamless transfer and storage of datasets via the Chemotion repository (Lütjohann 2015). (2) The laboratory repository Chemotion is the central collection point with long-term archival function for all recorded data of the syntheses and analyses that have been provided

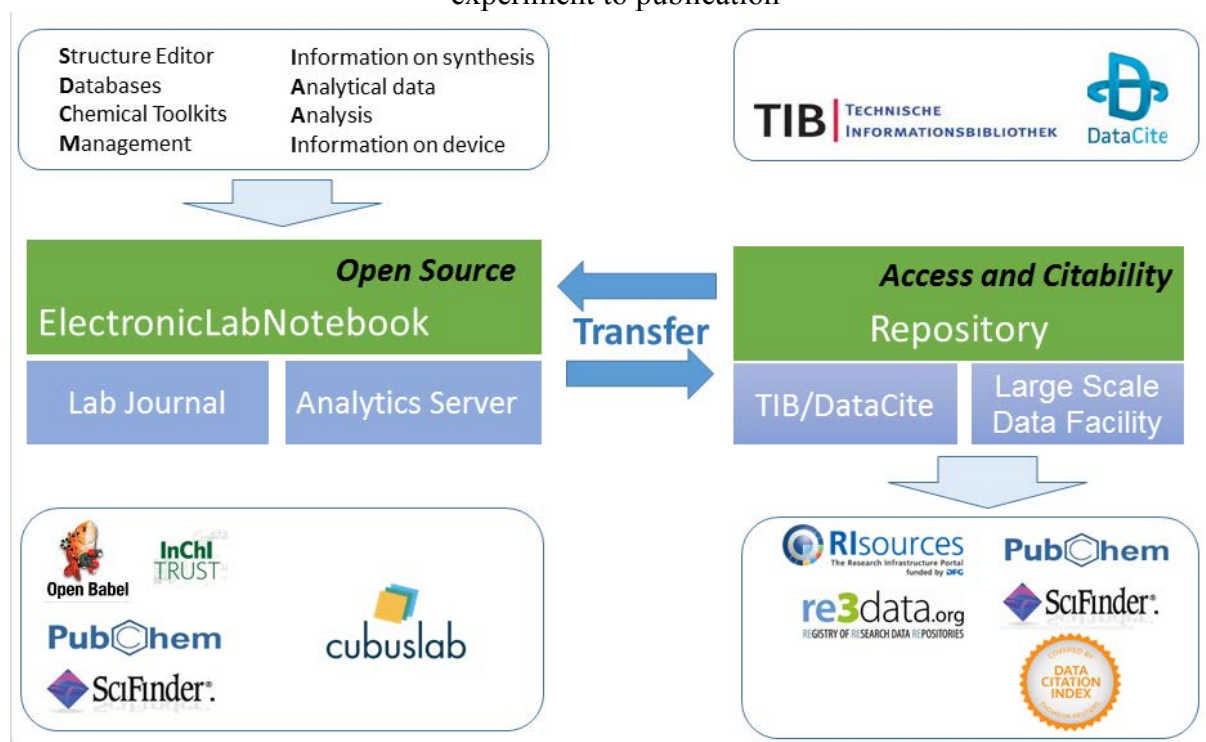
by single researchers (Jung 2016). All information that is available can be provided and shared with others via upload protocols for structures, analytics or descriptive resources. All developments of module 1 and 2 have been implemented as Open Source infrastructure to allow wide application and participation of other interested researchers.

RESULTS

Chemotion has become searchable as part of the registry of research data repositories re3data (Pampel 2013), a service of DataCite, listing about 1.500 research data repositories.

All compounds and datasets in Chemotion can be published and are automatically citable via DOI (Harvey 2015). The datasets are indexed in the subject specific databases PubChem and SciFinder, in the interdisciplinary Data Citation Index as well as other databases (Figure 1). We examined how the workflow for the citation of data sets can be improved to increase their visibility. Visibility and retrievability is an important incentive for researchers to enhance the culture of data publication and sharing. We show that improved data citation is well on the way to become an important aspect towards tracking full impact of research output.

Figure 1: Chemotion SMART LAB: organization and storage of research data from the single experiment to publication



REFERENCES

- Frey, J. G., & Bird, C. L. (2013). Cheminformatics and the Semantic Web: adding value with linked data and enhanced provenance. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 3(5), 465–481.
- Harvey, M. J., Mason, N. J., McLean, A., & Rzepa, H. S. (2015). Standards-based metadata procedures for retrieving data for display or mining utilizing persistent (data-DOI) identifiers. *Journal of Cheminformatics*, 7(1), 4.
- Jung, N., *Chemotion*. Retrieved March 16, 2016, from <http://www.chemotion.net/about>, re3data.org. <http://doi.org/10.17616/R34P5T>
- Lütjohann, D. S., Jung, N., & Bräse, S. (2015). Open source life science automation: Design of experiments and data acquisition via “dial-a-device”. *Chemometrics and Intelligent Laboratory Systems*, 144, 100–107.
- Nature Biotechnology (2009). Credit where credit is overdue. *Nature Biotechnology*, 27(7), 579.
- Pampel, H., Vierkant, P., Scholze, F., Bertelmann, R., Kindling, M., Klump, J., et al. (2013). Making Research Data Repositories Visible: The re3data.org Registry. *PLoS ONE*, 8(11), e78080



Changes in Scholars' Scientific Knowledge Production Shaped by Bibliometric Measures in Taiwan

Ming-Te Peng

*mpeng042@gold.ac.uk

Department of Sociology, Goldsmiths, University of London, New Cross, London, SE14 6NW, UK

ABSTRACT

Currently, bibliometrics have been playing an indispensable role in academic evaluation and resource distribution. In response to being measured, scholars might change their strategies to produce and deliver knowledge. Using the case study of biomedicine in Taiwan, my study shows that several strategies are utilised to maximise the 'performance' of scholars, including the model of cooperation for co-authors, the inclination for popular topics instead of a neglected branch and a decrease in translated works. This case study indicates that the application of bibliometric methods causes changes in researchers' motivations, perceptions and behaviour, interrupting the targets' activity patterns.

INTRODUCTION

Journal ranking systems are a quantitative order of academic international journals conducted by calculating the ratio of citation numbers of each academic journal. Nowadays the journal ranking system is regarded as a criteria to judge not only the value of an academic journal but also the quality of a research outcome (Burrows, 2012). Hence bibliometric indicators have become a major criteria to distribute research funding in many countries, known as performance-based research funding systems (Weingart, 2005).

In this case study, I explored the unintended consequence of the exercise of performance-based research funding systems in Taiwan and investigated if phenomena discovered in previous studies would be duplicated, like low risky conventional research and self-citations (Archambault & Larivière, 2009; Weingart, 2005).

METHODS

In this study I focuses on the field of biomedicine and all primary data were collected from three interviews that were conducted with open-ended questions. These three anonymous interviewees are professors in the biomedicine discipline in Taiwan. The open-ended interviews consisted of the same set of questions and allowed them to describe their personal experience and opinions. The open-ended questions concentrated on the researcher's action under the circumstances of being audited and their attitudes towards work that sits outside the bibliometric assessment.

RESULTS

After bibliometric measures were applied in performance-based research funding systems in Taiwan, some strategies have been adopted by researchers to behave like a productive and effective scholar as a kind of game. During my interviews the following patterns emerged.

First, one of the common tips to accumulate numerous publications in a short period is to 'team up'. To team up implies that researchers collaborate with their colleagues, while

members of the same group share each other's publications as co-authors. The process is outlined by the statement of one of my interviewees:

“Some of my colleagues choose to join in a ‘team’ and are listed as one of the co-authors in their respective studies to meet the criteria of published paper quantity. For example, in a team with five people, every member would publish one paper and lists all others as co-authors. Therefore, each of them will have five papers in the same time period.”

Because the frequency of being cited by other scholars is a significant indicator of quality papers, members of the same team may cite other member's previous articles as much as possible to raise the citation numbers. In other words, it can be regarded as a transformation of ‘self-citation’. In addition, the team may include clinical doctors, who are burdened with clinical, teaching and daily administrative work but also exposed to research evaluation. Hence, they are likely to cooperate with researchers by supplying clinical data and even sharing research funding as ‘contribution’, by which the doctor is able to be listed as a co-author.

To choose a low risky topic is another strategy. According to one interviewee's experience,

“Some of researchers like to do the same gene in different diseases or different organs in order to get more quick papers.”

Finally, activities that disperse scientific knowledge may be neglected because they are not identified as impact factors. For example, one interviewee expressed that he/she is inclined not to translate a scientific literature into Chinese because “it takes too much of my time that will reduce the time to analyse the experiments and write papers”. Another interviewee pointed out in his/her school there is a prize for excellent translated works, but the award has not been granted to anyone for several years because researchers are too busy to translate foreign works.

DISCUSSION

According to Espeland and Sauder's study of university ranking, the impact of public performance measures on academia can be partly analysed by Foucault's framework of discipline in terms of surveillance and normalisation (Foucault, 1979; Sauder & Espeland, 2009). In this case, the hierarchical publication system could be seen as a kind of panopticon, playing consistent, public and anonymous roles in monitoring and stratifying all academic members. Under bibliometric measures as surveillance, invisible intelligent efforts are entirely recorded with documentary accumulation, and the difference among researchers is subsumed into scales in an order. On the other hand, the funding formula play a key role in normalisation by defining the criteria of effective product and then rewarding excellent scientists. Taken together, the combination of the performance-based funding system and the bibliometric measure has modelled an environment with omnipresent social pressures. As one interviewee said, under this framework researchers “publish a paper like a xerox machine” and “lose the judgement about the true value of a publication”.

REFERENCES

Archambault, Éric, & Larivière, Vincent. (2009). History of the journal impact factor: Contingencies and consequences. *Scientometrics*, 79(3), 635-649.


Burrows, Roger. (2012). Living with the h-index? Metric assemblages in the contemporary academy. *Sociological Review*, 60(2), 355-372. doi: 10.1111/j.1467-954X.2012.02077.x

Foucault, Michel. (1979). *Discipline and punish : the birth of the prison / Michel Foucault ; translated from the French by Alan Sheridan: Harmondsworth : Penguin, 1979.*

Sauder, Michael, & Espeland, Wendy Nelson. (2009). The discipline of rankings: Tight coupling and organizational change. *American Sociological Review*, 74(1), 63-82.

Weingart, Peter. (2005, conf start date: 11-05-2003

cover date: 2005). Impact of bibliometrics upon the science system: Inadvertent consequences? (English).





Purpose-oriented metrics to assess researcher quality¹

Kedma Duarte^{*}, Rosina Weber^{**}, Roberto C. S. Pacheco^{*}

kedmaduarte@gmail.com, pacheco@egc.ufsc.br

^{*}Graduate Program in Knowledge Engineering and Management, Federal University of Santa Catarina, Brazil

rosina@drexel.edu

^{**}College of Computing and Informatics, Drexel University, USA

ABSTRACT

This paper contributes, illustrates, and evaluates a purpose-oriented method to assess quality of researchers. It is in selection processes such as recruitment that assessing the quality of researchers becomes necessary. Because quality is fitness for use, we contend that assessing the quality of researchers depends on references of quality imposed by the purpose of the selection process that motivates the need for assessment (henceforth target selection process). Our purpose-oriented method meets all requirements from The Leiden Manifesto for Research Metrics. We illustrate the method acquiring quality references from exemplars of researchers. These exemplars are presented as curriculum vitae of researchers (CVs) labeled as fit or unfit to target selection processes (i.e., binary quality assessment). We show two different target selection processes. The first considers fit researchers who are successful in collaborative endeavors, and thus are expected to succeed in collaborative research. The second considers fit researchers who are successful in solo endeavors, and thus are expected to succeed in solo research. We demonstrate that, on average, a classifier trained with data using quality references for a specific purpose (i.e., tailored to the peculiarities of their context) is more accurate than a purpose-independent classifier that does not consider the context of its target selection process.

INTRODUCTION

This paper introduces *purpose*-oriented methods that assess researcher quality and demonstrates they have the potential to be more accurate than methods that do not consider *purpose*. Assessing researcher quality is necessary in selection processes such as recruitment, promotion, and grant awarding decisions (Lane, 2010). Our method aligns with recommendations from international authorities in science & technology metrics (OECD, 2008) by including *purpose* in quality metrics because quality is *fitness-for-purpose* (Juran & Godfrey, 1999).

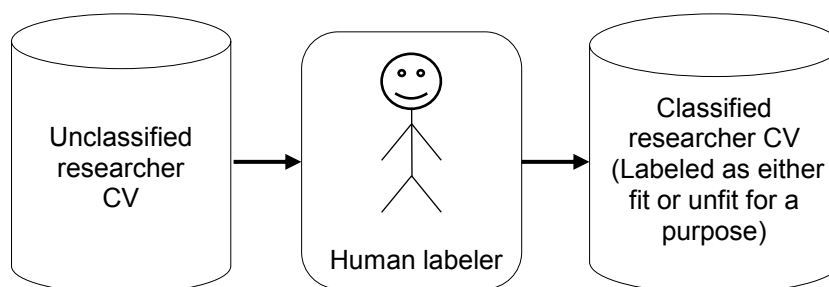
The Leiden Manifesto (Hicks & Wouters, 2015) is a set of best practices for applying metrics for assessing research quality. The manifesto suggests attention to transparency, flexibility, and contextual elements important to the *purpose* of the assessment.

Our approach integrates *purpose* with information about the selection process that needs to assess researcher quality. For example, in a recruitment process, the quality assessment is

¹ First author is supported by Brazilian's Goiás Research Foundation (FAPEG) and University of the State of Goiás (UEG) under agreement number 201310267000099. Authors thank the STELA Institute, particularly Rudger Taxweiler's help with collecting data.

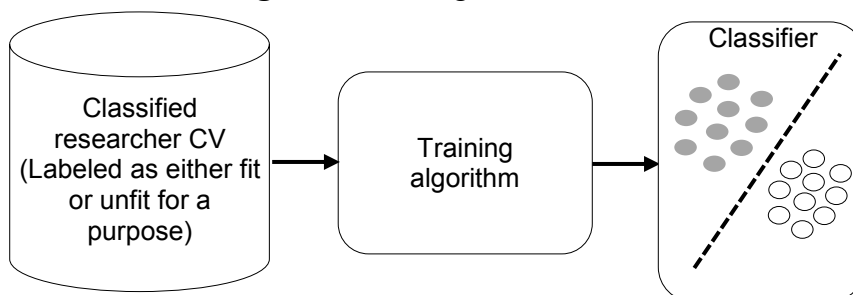
used to determine how fit different applicants are to fill a job, guiding the decision of whom to hire. The information may describe explicit parameters such as relevant outputs (e.g., articles or books) or as examples of curricula vitae (CV) representing candidates that would be fit or unfit for a job. We investigate how considering or not *purpose* influences the accuracy of the assessment; that is, does the method correctly classify fit and unfit candidates?

Figure 1. Humans can label instances to train classifiers

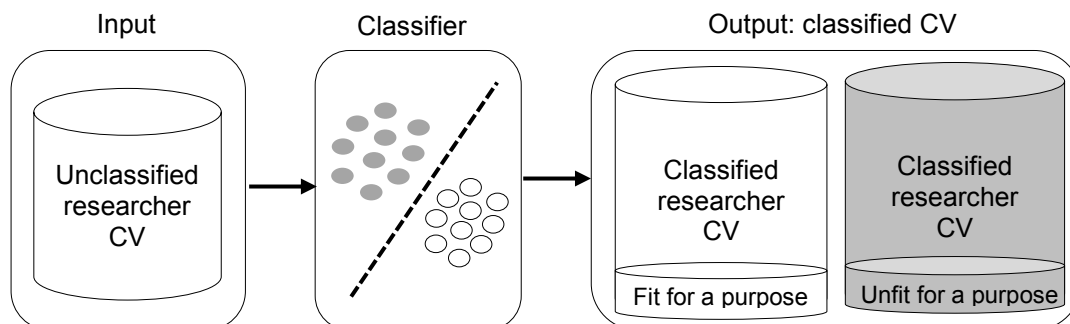


Our approach is based on machine learning, which is concerned with training algorithms that learn knowledge from data. The training algorithm produces a classifier, which is a data structure that can classify, for example, a CV of a researcher as fit or unfit for a *purpose* (i.e., job).

Figure 2. Training the classifier



Humans can label data instances (e.g., curricula vitae) to train classifiers (Figure 1). The resulting classifier exhibits the same behavior as the relation between the instances and the labeled instances used for training. In the example of training data labeled as fit and unfit, the classifier (Figure 2) is expected to classify new, previously unseen CV, like human labelers (Figure 3).

Figure 3. Using the classifier to assess researcher quality for a purpose

Classifying candidate CV as fit or unfit is how the method assesses researcher quality. In this example, the job is the purpose. This is the basic principle utilized in many big data applications. The proposed classifier is based on weight learning and case-based reasoning (Richter & Weber, 2013).

PURPOSE-ORIENTED OR PURPOSE-INDEPENDENT ASSESSMENT?

In this section, we demonstrate the hypothesis, “*A purpose-oriented method to assess researcher quality is more accurate than a purpose-independent method*”. In this work, we adopted the method above described with purposes being two job openings: one for a collaborative researcher (*collaboration job*), the other for a researcher to work alone (*solo job*). We tested this hypothesis using CV data from the Brazilian Lattes database (Pacheco et al., 2006), a high quality profiling system (Lane, 2010).

The purpose is incorporated into the method via CV labeled by the authors as representing example candidates that are fit and unfit for each of the two jobs. These parameters were then used to label the unclassified CV data. We refer to this set of classified CV as fit or unfit as *ground truth*, which is used as reference of accuracy. We define accuracy as the percentage of correct classifications to the total number of classified CV.

This ground truth dataset is also used to train weights for the three classifiers: the two purpose-oriented classifiers, and the purpose-independent classifier. To deploy this approach in a real-world quality assessment, a small set of CV would need to be labeled with examples of fit and unfit candidates for training.

There are 28 applicants for the collaboration job, and 29 for the solo job. We train two purpose-oriented classifiers (collaboration and solo jobs). We trained one purpose-independent classifier (i.e., does not distinguish jobs). We compared the accuracy between the two purpose-oriented (POC1, POC2) against the purpose-independent classifier (PIC).

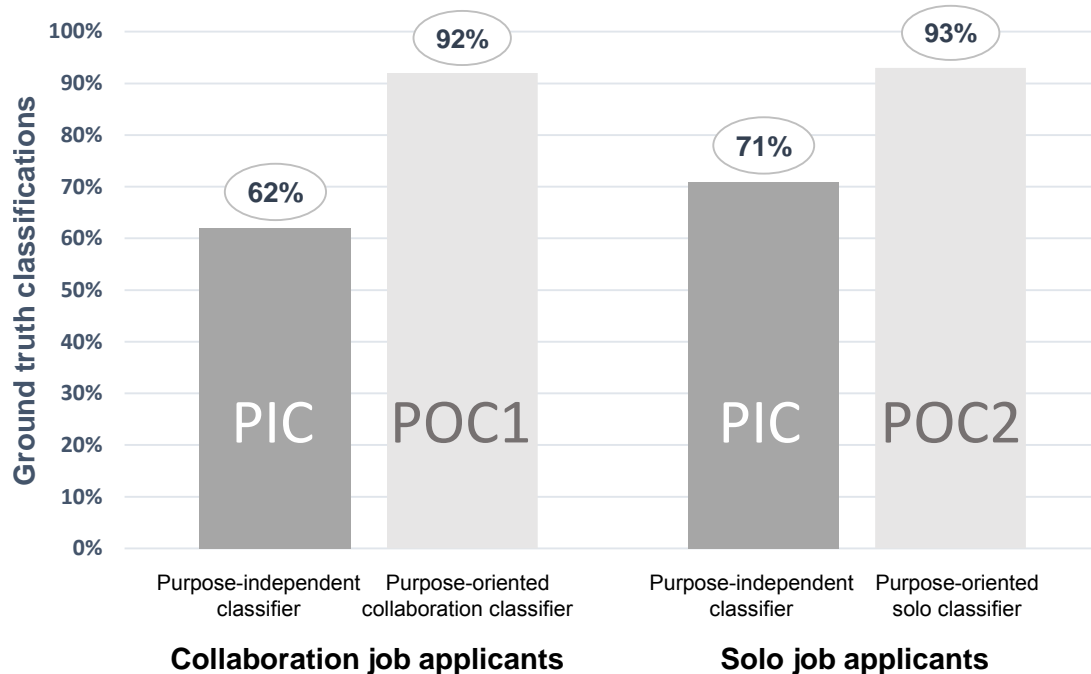
Figure 4. Accuracy by classifier and by job

Figure 4 shows accuracy for all classifiers. The purpose independent classifier correctly classified 62% of the applicants for the collaboration job, and 71% for the solo job. The purpose-oriented classifiers correctly classified 92% and 93%, respectively. These results support our hypothesis that purpose-oriented classifiers are more accurate than purpose independent. These levels of accuracy of the purpose-independent classifier would falsely consider unfit five and four applicants, respectively, that are actually fit for the collaboration and solo jobs. The performance of the purpose-oriented methods would have falsely labeled only one applicant in each of the jobs.

CONCLUSIONS

We introduced, described, and validated a purpose-oriented method to assess researcher quality. Our approach is aligned with recommendations from the OECD (2008) to incorporate purpose based on the notion that quality is *fitness-for-purpose* (Juran & Godfrey, 1999). Our approach integrates purpose of assessment through examples of fit and unfit researchers. It implements the second principle of the *Leiden Manifesto* (Hicks & Wouters, 2015), which argues that performance should consider contextual aspects, which we represent through examples of fit and unfit researcher CV.

REFERENCES

- Hicks, D. & Wouters, P. (2015). The Leiden Manifesto for research metrics. *Nature*, 520, 429–431.
- Juran, J. M., & Godfrey, A. B. (1999). *Juran's Quality Handbook*. New York: McGraw-Hill. 5th Edition.
- Lane, J. (2010). Let's make science metrics more scientific. *Nature*, 464, 488–489.
- OECD (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide*. Paris: OECD Publications.
- Pacheco, R. C. S., Kern, V. M., Salm Jr, J. F., Packer, A. L., Murasaki, R., Amaral, L., Santos, L.D., Cabezas B., A. R. (2006). Toward CERIF-ScienTI cooperation and interoperability. In: A.G.S. Asserson & E. J. Simons (Eds.) 8th International Conference on Current Research Information Systems (p.179–188). Leuven: Leuven University Press.
- Richter, M. M., & Weber, R. O. (2013). *Case-Based Reasoning: a Textbook*. Berlin: Springer-Verlag.



On the relationship between research topics and scientific impact: a study of edible animal research

Lourdes Castelló-Cogollos^{*}, Rafael Aleixandre-Benavent^{*,**}, Pablo d'Este^{**} and Ismael Ràfols^{**,***}

^{*} lourdes.castello@uv.es, rafael.aleixandre@uv.es

Unitat d'Informació i Investigació Social i Sanitària, UISYS(CSIC-UV)
Dept. de Sociologia i Antropologia Social, Universitat de València, València, 46003 (Spain)

^{**} pabdescu@ingenio.upv.es

Ingenio (CSIC-UPV), Universitat Politècnica de València, València, 46022 (Spain)

^{***} i.rafols@ingenio.upv.es

SPRU (Science Policy Research Unit), University of Sussex, Brighton, BN1 9SL (England)
Observatoire des Sciences et Techniques (OST-HCERES), Paris (France)

ABSTRACT

Finding appropriate units for comparison has long been a central and highly contentious issues in evaluative bibliometrics (Ophoff and Leydesdorff, 2010; Waltman et al., 2011). The conventional approach has been to make comparisons within a given field of research – which poses a number of challenges such as field size (Zitt et al. 2005; Adams et al., 2008) and field delineation (Rafols and Leydesdorff, 2009; Boyack and Klavans, 2015). Van Eck et al. (2013) showed recently that conventional field-size normalisation based on Web of Science Categories is problematic in medical research, since within a given field it tends to favour basic over clinical research.

The hypothesis of this article is different topics have different scientific impact as a result of their different research size. This means that the most highly cited articles on an issue that is small because of its local specificity are far less cited than cited articles on an issue that captures many scientific and economic resources.

In this study we investigate research on different edible animal in the WoS Category of “Agriculture Dairy Animal Science”, which concerns the rearing of animals as food. The question is whether animals such as cows and sheep, which are of interest (because they are eaten) in many countries have higher visibility than articles concerned with animals such as rabbits, ducks or buffalos, which are only of interest in a limited number of countries.

The first hypothesis is thus that cows and sheep will have a relatively larger proportion of articles in higher quartile journals (Q1), in the top 10% cited papers, and mean number of citations, than articles on rabbits and buffaloes.

The second hypothesis is that part of the low impact of some countries (such as Spain or India) is due to the fact that they work on less popular topics (such as rabbits or buffalos) and the high impact of central countries (such as the US or the Netherlands) is partly due to the focus on mainstream topics.

However, we do not expect either hypothesis 1 or 2 to be the only relevant factor, because there may be indeed differences in the research quality (i.e. scientific interest as perceived by scientific peers) of research carried out in certain animals and countries.

In this poster we will present the analyses of regression that will tease out the different influence of topic (animal) size, topic and country. These analyses will allow to test to which extent hypotheses 1 and 2 are supported.

We believe that the results can be potentially relevant for evaluative bibliometrics. If hypothesis 1 is confirmed, evaluations may need to take into account whether the topics examined have a local nature and thus have size constraints.

Table 1. Differences in topic size, citations and journal quartile (2006-2015).

Animals	Documents	Cits/Doc	Q1 (%)
Pig	6292	6.98	49.08
Cow	5670	7.93	57.13
Sheep	4083	5.83	27.95
Chicken	3969	6.40	44.27
Goat	2632	5.53	20.90
Buffalo	1427	2.23	7.50
Horse	894	6.16	47.54
Rabbit	800	4.30	21.13
Turkey	462	5.11	48.48
Bull	400	5.97	32.00
Duck	321	3.17	48.91
Quail	268	4.11	36.94
Camel	232	4.84	18.10
Yak	145	2.50	17.93
Deer	105	3.67	20.00

REFERENCES

Adams, J., Gurney, K., & Jackson, L. (2008). Calibrating the zoom—A test of Zitt's hypothesis. *Scientometrics*, 75(1), 81-95.

Klavans, R., & Boyack, K. W. (2015). Which type of citation analysis generates the most accurate taxonomy of scientific and technical knowledge?. arXiv preprint arXiv:1511.05078.

Opthof, T., & Leydesdorff, L. (2010). Caveats for the journal and field normalizations in the CWTS ("Leiden") evaluations of research performance. *Journal of Informetrics*, 4(3), 423-430.

Rafols, I., & Leydesdorff, L. (2009). Content-based and algorithmic classifications of journals: Perspectives on the dynamics of scientific communication and indexer effects. *Journal of the American Society for Information Science and Technology*, 60(9), 1823-1835.

Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, 5(1), 37-47.

Waltman, L., & van Eck, N. J. (2013). Source normalized indicators of citation impact: An overview of different approaches and an empirical comparison. *Scientometrics*, 96(3), 699-716.

Zitt, M., Ramanana-Rahary, S., & Bassecoulard, E. (2005). Relativity of citation performance and excellence measures: From cross-field to cross-scale effects of field-normalisation. *Scientometrics*, 63(2), 373-401.



Evaluation of grants schemes in the context of the national research system based on the publication count and citation data: the grants of the Latvian Council of Science ¹

Arnīs Kokorevics*

* arnis.kokorevics@lzp.gov.lv

Latvian Council of Science, Akademijas sq. 1, Riga, LV1050 (Latvia)

ABSTRACT

The study analyses the Fundamental and Applied Research projects of the Latvian Council of Science (LCS FARP). The number of publications and their citation analysis are compared with those of all the national output during the same period. This work testifies such an approach when the output (number of publications) and quality of such an output (citation characteristics) for separate national grants schemes is compared with the common national output if the grant schemes cover all branches of sciences. Therefore, it will be able to compare. The influence of the international cooperation publications as hyperauthorship publications must be taken into account in such a comparison. The publications of LCS FARP show ~30% lower citation characteristics during 2010-2012 and ~17% lower during 2013 than all Latvia's publications in total. If the hyperauthorship publications with ≥ 50 authors (0.5-0.75 per year) are excluded from the comparison, the differences between both the groups tend to smooth out: from ~25% in 2010 to 9% in 2013.

INTRODUCTION

Alongside the evaluation of the scientific quality and impact of institutions and countries, urgent is also the evaluation of R&D&I policy instruments and funding sources. It allows an understanding of its role and contribution in the national research system, comparing policy instruments and funding schemes, and if necessary, advocating it. Both policy makers and the institutions involved into the management of research are interested in receiving such feedback as soon as possible, during or just after the ending of projects.

The Fundamental and Applied Research projects of the Latvian Council of Science (LCS FARP) represent a kind of grants, which are devoted to creating new knowledge and technological thoughts with the aim of promoting high quality research. As grants, the themes of which are initiated by researchers, they occupy their own place in the national research system. The study by Kokorevics, Kunda & Bundule (2014) analyses the return of the LCS FARP realized during 2009-2012 (thematic projects, as usually realized by 1 institution) and 2010-2013 (cooperation projects, realized by 3-4 national partners). Alongside the counting of different types of the outputs mentioned in the reports and surveys of grant leaders, this study includes also the publications count and their citation analysis. Taking into account the fact that LCS FARP covers all branches of sciences (including social sciences, arts and humanities), the number of publications and characteristics of citation are compared with

¹ This work was supported by the Latvian Council of Science, Riga, Latvia.

those of all the national output during the same period, because such a comparison follows the principle “compare similar with similar”. This work testifies such an approach when the publication output of a separate kind of grants is compared with all the national output to evaluate the comparable scientific quality in the context of the national research funding landscape.

METHODS AND DATA

Thomson Reuters Web of Science™ Core Collection (WoSCC) is used as the data source and analysis facilities offered by this database are utilized for the study.

To perform the predicted task – comparison of the output of a certain group of research grants – LCS FARP with the national totality of publications, two questions are relevant:

- 1) Which scope (types) of publications will be appropriate to be chosen for comparison?
- 2) How to recognize the publications, which represent the output of a certain group of research grants (in this case – LCS FARP)?

A preliminary comparison (17.05.2014) of Latvia’s publications (records having the term “Latvia*” in the *Address* field) from year 2011 shows that 593 publications of the type *Article*, *Review* and *Letter* in SCI-EXPANDED, SSCI, A&HCI represent only 47.2% of all Latvia’s publications in WoSCC (1256), but receive 95.5% of citations (2887 from 3023). The proportion of self-citation for this group is only 1.56%. For further analysis, only these types of publications are chosen. To estimate the effect of the so-called hyperauthorship publications (Cronin, 2001), the publications with 50 or more authors were detected and the characteristics of citation were calculated to the groups of publications with and without such hyperauthorship publications.

A modified approach used by Belter (2013) has been utilized to recognize the publications belonging to LCS FARP: the publications mentioning the financial support of LCS FARP in their acknowledgements (in records of WoSCC) and the publications stated in yearly reports of the projects. It should be noted that during 2009-2013, LCS did not require an obligatory acknowledgement to the offered financial support and did not suggest the unified form of such an acknowledgement. Therefore, the acknowledgement will not be sufficient to gather the publications of LCS FARP. The yearly reports (submitted in the next month after the completion of the period) included also the manuscripts submitted to press. The corresponding publications indexed by WoSCC were recognized manually. Nevertheless, the utilization of these two mutually complementary procedures did not allow detecting all the corresponding publications and will lead to an underestimating of the output of LCS FARP.

All the national output and the output of LCS FARP as the number of publications are counted for each year within the period of 2009-2013, and their citation characteristics as average citations per publication were calculated for the publications of each year within the year of publication and the following years. The data used in Kokorevics, Kunda & Bundule (2014) (calculated on 08.07.2014) were updated (06.03.2016) and the citation analysis is expanded from 2009-2013 to 2009-2015. These data are deposited and available as Open Data (Kokorevics, 2016). The projects statistics (<http://www.lzp.gov.lv/>) and official statistical data (<http://data.csb.gov.lv/>) for the utilized financial and human resources are used additionally to scientometrics data.

RESULTS AND DISCUSSION

For the evaluation of the contribution of LCS FARP in the context of the national research system, the three topics will be discussed:

- 1) the quantity of the output (in terms of the numbers of publications) and its comparison with the input (in terms of the utilized financial and human resources);
- 2) the quality (in terms of citations);
- 3) the stability of the quality (in terms of citations).

From methodological aspects, an interesting topic is the appropriateness of such proposed approach to evaluate the grants schemes or financing instruments during or just after the ending of activities based on comparison with the national totality of publications.

The publications referable to LCS FARP represent 31.7% of all the national output of the same type during 2009-2013 and have a maximum during 2010-2012 (Table 1). While all the national output exhibits an essential rise in 2011 (from 422 in 2010 to 595), the LCS FARP exhibits an increase in the period of 2009-2011, the same level of output in 2012 comparing to 2011 and a drastic reduction in 2013. It would be explained by the project realization cycle: the thematic projects (which receive 67-70% financing) started in 2009 and ended in 2012. Therefore, the results accumulated in these research activities resulted in publications in 2011-2012. The new cycle of thematic projects started in 2013 did not result in a great amount of publications immediately. To estimate the contribution of LCS FARP, it is essential to compare it with the utilized financial resources and human resources involved. The proportion of the LCS FARP financing from all the national financial support for research decreased from 4.4% to 2.3% during 2009-2013 (from the state financial support for research, it varied between minimum 9.5% in 2012 to maximum 11.6% in 2010). At the same time, it will not be excluded that some part of publications is based on the research activities, which utilize also other financial resources additionally to the LCS FARP funding. During 2010-2012, when a permanent number of thematic and cooperation projects had been realized its involved 14.7% of researchers in Latvia (head counting). Therefore, a substantial part of the most significant output of Latvia's scientific publications (Articles, Reviews and Letters indexed by WoSCC) is linked to the research activities realized within LCS FARP, and its part exceeds several times the proportion of the utilized financial resources. Also, the proportion of the produced publications exceeds at least twice the proportion of the involved Latvian researchers.

Table 1. Number of Latvia's and LCS FARP publications in 2009-2013 (*Articles, Reviews, Letters* indexed by SCI-EXPANDED, SSCI, A&HCI).

Year	Latvia's publications		LCS FARP publications		Proportion LCS FARP vs Latvia (%)
	All	≥50 authors	All	≥50 authors	
2009	436	1	127	0	29.1
2010	422	3	153	1	36.3
2011	595	3	226	1	38.0
2012	601	3	216	1	35.9
2013	640	4	132	0	20.6
Total	2694	14	854	3	31.7

The quality of publications is evaluated in terms of citation, received after the publication. In an earlier study (Kokorevics, Kunda & Bundule (2014)), only the citation data for 2009-2013 was available. In this study, the data is prolonged to 2015. Therefore, we can utilize the citation characteristics for the issuing year and the next four years for the publications issued in 2009-2011 and by 1 and 2 years shorter periods for publications of 2012 and 2013,

respectively (Figure 1). Although the publications of 2009 of LCS FARP show higher (~20%) citation characteristics in subsequent years than all Latvia's publications in total, for publications in the next years, the situation is opposite: ~30% lower citation characteristics for publications of 2010-2012 and ~17% lower for publications of 2013.

Therefore, the output of LCS FARP seems to be with lower quality than all Latvia's publications. Which kinds of research projects produce the best publications in Latvia? The available data did not allow evaluating the output of other funding instruments in Latvia during 2009-2013, because the acknowledgements praxis in different types of projects is not uniform but the report data of other funding institutions are not available. There was a possibility to evaluate the influence of hyperauthorship publications with ≥ 50 authors. If such publications (0.2-0.7% per year of all Latvia's publications) are excluded from the comparison, no changes are evaluated for publications of the issuing year 2009 and 2011, while for other years, differences between both the groups tend to smooth out, especially for publications of 2012. This confirms the conclusions of Allik (2013) and Must (2013) that such hyperauthorship publications can substantially affect the scientometrics characteristic in the case of countries having a small or moderate output of publications.

Figure 1. Average citations per publication of the publications of 2009-2013 depending of the years after publication (publication year - 0, next years - 1-4). Comparison of citations of all Latvian and LCS FARP publications with and without publications with ≥ 50 authors (Cor.).

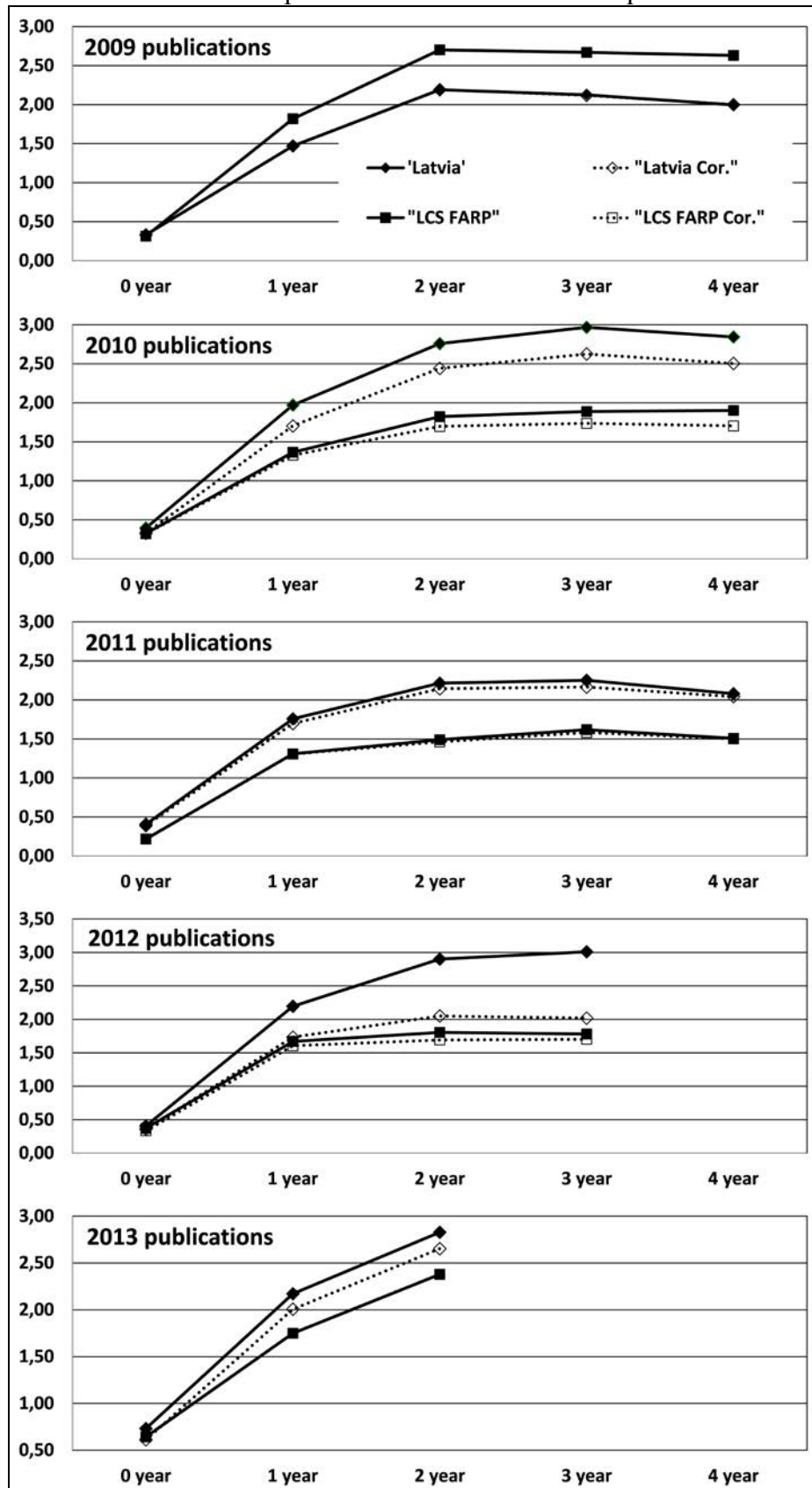
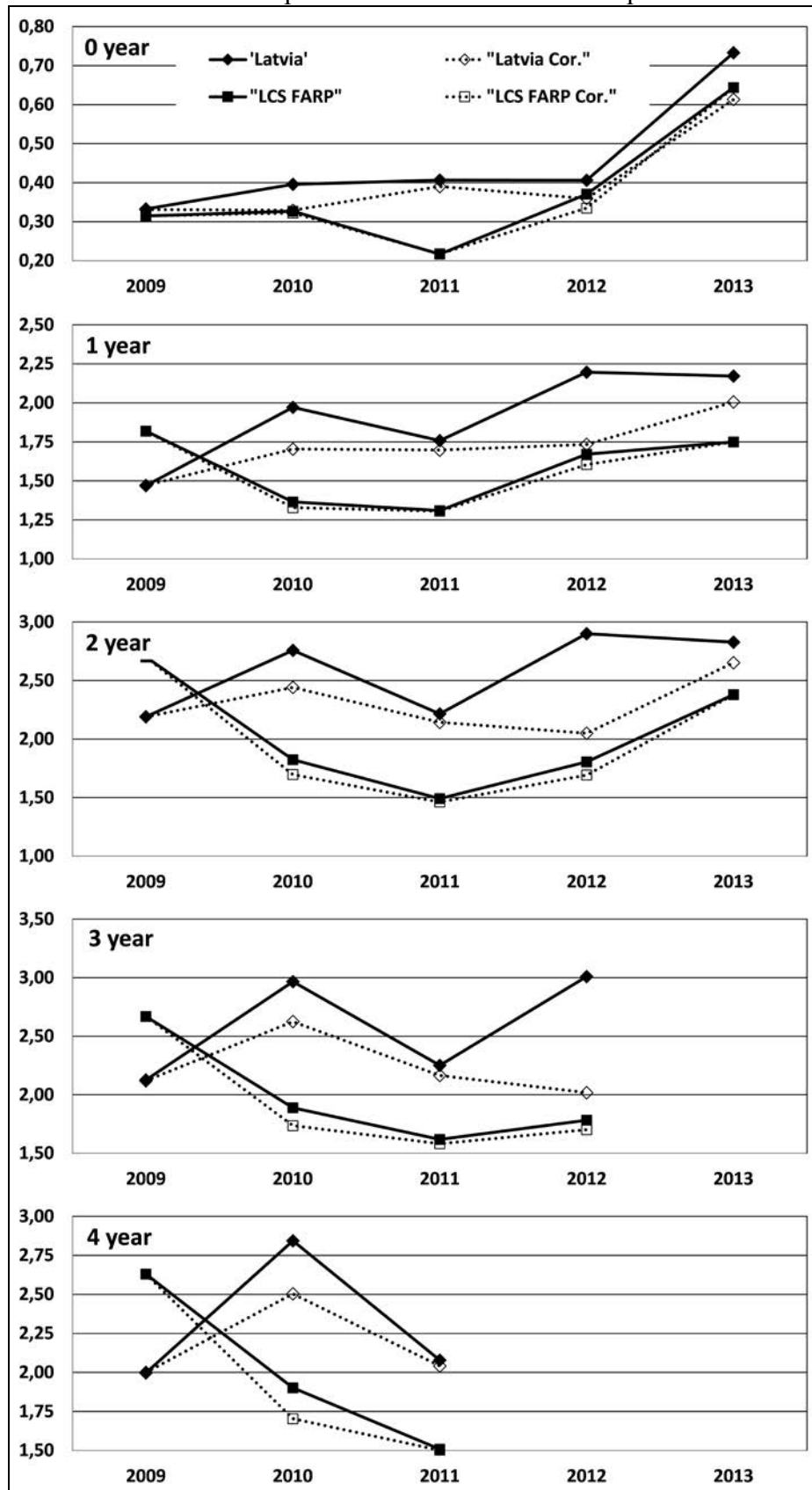


Figure 2. Average citations per publication during the publication year (0 year) and the next years (1-4 years) depending of the publication year 2009-2013. Comparison of citations of all Latvian and LCS FARP publications with and without publications with ≥ 50 authors (Cor.).



Researchers recognize that the publications indexed in WoSCC and Scopus will promote their career, increase the possibility to gain new projects and encourage the evaluation of their institutions. There is also some pressure today from the administration of institutions and research funding organizations for such level publications. Therefore, it is possible that research groups prefer to increase the quantity (the number of publications) instead of quality, which will be expressed in citation characteristics. LCS FARP and all Latvian output exhibit a rise in the number of publications within 2009-2013. Is the quality stable, increasing or decreasing? The publications issued in 2009-2013 are compared in the issuing year (0 year) and during the next years (1-4 years) (Figure 2). The publications of 2011 seem to exhibit a weaker citation than the publications of the previous year 2010 and the subsequent years 2012-2013. If the hyperauthorship publications with ≥ 50 authors are excluded from the comparison, the differences smooth out both between the LCS FARP and all Latvia's publications and between different publication years. Therefore, the observed variation in citations characteristics did not allow concluding about the existence of any tendencies that the quality of publications tended to change during this period.

Although this study prolongs the citation analysis by two years (2014-2015), the conclusions regarding the proposed three topics in this study remain the same as in the initial study (Kokorevics, Kunda & Bundule (2014)). Therefore, it will be able to compare the output (number of publications) and quality of such an output (citation characteristics) for separate national grants schemes with the common national output if the grant schemes cover all branches of sciences. The possibility to realize the comparison during or just after the ending of projects is especially valuable. The influence of the hyperauthorship publications, which are usually international cooperation publications, must be taken into account in such a comparison. More reliable data for the evaluation of national scale grants schemes will be achieved by excluding such publications from comparison.

REFERENCES

- Allik, J. (2013). Factors affecting bibliometric indicators of scientific quality. *Trames*, 2013, 17, 199-214.
- Belter, C. W. (2013). A bibliometric analysis of NOAA's Office of Ocean Exploration and Research. *Scientometrics*, 95, 629-644.
- Cronin, B. (2001). Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices? *Journal of the American Society for Information Science and Technology*, 52, 558-569.
- Kokorevics, A., Kunda, I., & Bundule, M. (2014). Return of the basic and applied research projects of the Latvian Council of Science within the context of socioeconomic categories. Riga: Latvian Council of Science. Retrieved March 1, 2016 from http://www.lzp.gov.lv/parskati/LZP-petijums-LZP_Projektu_atdeve_2014.pdf (in Latvian).
- Kokorevics, A. (2016). Data set of Latvia's and LSC FARP publications in 2009-2013 and their citation during 2009-2015 based on WoSCC. Retrieved July 10, 2016 from <https://www.researchgate.net/publication/299468499> (in Latvian).
- Must, Ü. (2013). The impact of multi-authored papers: the case of a small country. *Collnet Journal of Scientometrics and Information Management*, 8, 41-47.



New approaches to monitor and evaluate Science, Technology and Innovation in health: a pilot study on the Zika virus

Paula Xavier dos Santos^{*}, Rebeca Buzzo Feltrin^{*}, Bruna de Paula Fonseca e Fonseca^{**},
Ricardo Barros Sampaio^{***}, Juliana Gonçalves Reis^{*}, Jorge Barreto^{***}, Fatima Martins^{****},
Maurício Barreto^{*****}, Nísia Trindade Lima^{*}

^{*} paula.xavier@fiocruz.br, rebeca.feltrin@fiocruz, juliana.goncalves@fiocruz.br, nisia.lima@fiocruz.br
Vice-Presidency of Education, Information and Communication (VPEIC). Av. Brasil, 4365 - Manguinhos, Rio de Janeiro - CEP:21040-900. Rio de Janeiro, Brazil.

^{**} bfonseca@cdfs.fiocruz.com
Center for Technological Development in Health, Oswaldo Cruz Foundation, Av. Brasil 4036, Rio de Janeiro, 21040-361 (Brazil)

^{***} ricardosampaio@fiocruz.br, jorgeomaia@hotmail.com
Diretoria Regional de Brasília, Oswaldo Cruz Foundation, Av L3 Norte, s/n, Campus Universitário Darcy Ribeiro, Gleba A, Brasília, 70910-900 (Brazil)

^{****} fatima.martins@icict.fiocruz.br Av. Brasil, 4365 - Manguinhos, Rio de Janeiro - RJ, 21040-360 (Brazil).
Library Network, Icict, Fiocruz.

^{*****} mauricio.barreto@bahia.fiocruz.br
Senior Researcher, Fiocruz-Bahia and Professor of Epidemiology, UFBA - Epidemiologia, Saúde Coletiva, Social Epidemiology, Public Health, Impact Evaluation

ABSTRACT

The tools for evaluation of research and technological development (R&D) activities have many limitations related to the specific dynamics of knowledge production and socioeconomic realities of the peripheral countries. Among the consequences of these limitations are the formulation of policies based on issues that are external to these countries' realities and the relentless pursuit of the improvement on the position of institutions and countries in international research rankings, rather than of changes according to local demands. This is even more critical in situations that require fast and reliable government interventions, as in the case of the Zika virus epidemic.

Given this context, the Oswaldo Cruz Foundation (Fiocruz), has been leading the implementation of the Observatory in Science, Technology and Innovation in Health (STI in Health). The Observatory is a tool to support the management of STI activities in health for the formulation of institutional policies based on indicators, metrics and, principally, qualitative analysis. For this, it should adopt new approaches to monitor and evaluate research and technological development, designed according to the specificities, diversity and complexity of the dynamics of knowledge production in health, being able to reveal and value the results of these activities and their impact and benefits to society. The approaches developed in this project will be implemented in a pilot study on the Zika virus outbreak, in order to monitor and produce qualified information for political decision-making on the subject. Therefore, it should carry out i) the systematic monitoring of scientific production (articles and research projects) and ii) the mapping and analysis of collaborative networks of

researchers and research institutions in order to promote the use of scientific evidence in policy-making related to the Zika epidemic.

BACKGROUND

The tools for evaluation of research and technological development (R&D) activities have many limitations related to the specific dynamics of knowledge production and socioeconomic realities of the peripheral countries. Among the consequences of these limitations are the formulation of policies based on issues that are external to these countries' realities and the relentless pursuit of the improvement on the position of institutions and countries in international research rankings, rather than of changes according to local demands. This is even more critical in situations that require fast and reliable government interventions, as in the case of the Zika virus epidemic.

The Zika epidemic is now priority agenda of R&D in key health institutions in Brazil and the world (WHO, 2016; Fiocruz, 2016). In the field of information and communication in health, we can be experiencing an "information epidemic" that results in a paradox in which too much information eventually generate misinformation about what is relevant and reliable, both for citizens in general, and for everyone involved in the management of the public health system - health professionals, managers and the scientific community.

Given this context, the Oswaldo Cruz Foundation (Fiocruz), the most prominent public health institution in Latin America, has been leading since 2013, the implementation of the Observatory in Science, Technology and Innovation in Health (STI in Health). The Observatory is a tool to support the management of STI activities in health for the formulation of institutional policies based on indicators, metrics and, principally, qualitative analysis. For this, it should adopt new approaches to monitor and evaluate research and technological development, designed according to the specificities, diversity and complexity of the dynamics of knowledge production in health, being able to reveal and value the results of these activities and their impact and benefits to society.

The approaches developed in this project will be implemented in a pilot study on the Zika virus outbreak, in order to monitor and produce qualified information for political decision-making on the subject. Therefore, it should carry out i) the systematic monitoring of scientific production (articles and research projects) and ii) the mapping and analysis of collaborative networks of researchers and research institutions in order to promote the use of scientific evidence in policy-making related to the Zika epidemic.

OBJECTIVE

General objective

Theoretical design and implementation of a new approach to monitor and evaluate research activities and technological development through the integration of quantitative indicators and qualitative studies that seek, with the support of scientific evidence, the formulation of public policies for confronting the Zika epidemic.

Specific objectives

- i. Support the development of qualitative approaches that promote the use of scientific evidence for decision-making and formulation of public health policies;
- ii. Support decision-making and policy-making, especially in crisis situations such as the case of the Zika epidemic, through the use of scientific evidence;
- iii. Monitor in a dynamic and ongoing way the evolution of scientific literature, including papers and research projects related to the Zika virus;
- iv. Map the collaborative networks of researchers and institutions to develop Zika virus-related research.

METHOD

i) Mapping of collaborative networks of researchers and institutions dedicated to research related to the Zika virus. Scientific collaboration networks will be built based on co-authorship data retrieved from: i) scientific papers included in the main databases of specialized national and international data in health; and ii) scientific projects financed by the main agencies of Brazilian and international development.

ii) Systematic and dynamic monitoring of scientific production (publications and projects) based on the systematic review method. Develop a synthesis of scientific studies that address questions previously formulated by managers, in order to identify, select and critically appraise relevant research, as well as collect, analyze and synthesize data from this research (Brazil, 2014).

Expected preliminary findings

- i) Scientific evidence mapped from the published scientific papers and research projects submitted in Brazil in the years 2014-2016;
- ii) Mapping of collaborative networks between researchers and institutions based on the co-authorship relations identified in publications.

Discussion and conclusion

The theoretical approach and methods developed in this study will help the decision-making process through the surveying, organization, critical evaluation and production of qualified information, targeted to respond to the problems and needs of the processes of public policy formulation and implementation. The coordination between scientific production and information, seeking the quality of decision-making processes and promoting agility and reliability in decision-making in times of crisis provides an opportunity to think about the translation of knowledge and the application of new approaches and mechanisms to overcome the gaps of science so that R&D can reach the Brazilian population.

REFERENCES

Bornmann, L. (2013). What Is Societal Impact of Research and How Can It Be Assessed?: a literature survey. *Journal of the American Society for Information Science and Technology*, 64, 217–233.

Brazil. (2014). Ministry of Health. Evidence Synthesis for Health Policy: stimulating the use of scientific evidence in decision-making. Brasília: Ministry of Health.

FIOCRUZ. Fundação Oswaldo Cruz (2016). Fiocruz assina acordo de cooperação global contra zika. Matéria publicada no Portal da Fiocruz, 12/02/2016. Disponível em: <<http://portal.fiocruz.br/pt-br/content/fiocruz-assina-acordo-de-cooperacao-global-contra-zika>>

WHO. World Health Organization. (2016). Zika Strategic Response Framework & Joint Operations Plan.



Issues relating to a Brazilian model of graduate courses evaluation: the CAPES system

Michely Jabala Mamede Vogel^{*}, Nair Yumiko Kobashi^{**}

^{*} michelyvogel@id.uff.br

Instituto de Artes e Comunicação Social, Universidade Federal Fluminense, Rua Lara Vilela, 126, Niterói-RJ, 24210-590 (Brazil)

^{**} nairkobashi@gmail.com

Escola de Comunicações e Artes, Universidade de São Paulo, Av. Prof. Lúcio Martins Rodrigues, 441. São Paulo-SP, 05616-010 (Brazil)

ABSTRACT

Presentation and discussion of an evaluation model adopted in Brazil – the evaluation system of CAPES (Higher Education Personnel Training Coordination). Concepts related to the evaluation society proposed by Dahler-Larsen (2011) and the weight of the scientific reputation of researchers and research institutions in the context of academic competition, proposed by Bourdieu (1983; 2004) are the theoretical framework of this research. As final considerations, problems resulting from the use of this evaluation model are indicated.

INTRODUCTION

Scientific research activities are submitted, worldwide, according to universal parameters of evaluation, based on indicators that are supposed to be objective. According to Dahler-Larsen (2011), the evaluation process is almost sacred. However, scientific research, as any social activity, is structured according to the weight of the scientific reputation of the researchers or institutional agents that act over it (Bourdieu, 1983; 2004).

The evaluation process requires “the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about it, improve program effectiveness, or inform decisions about future programming” (Patton, 1997 as cited in Dahler-Larsen, 2011).

An important question to keep in mind is that an evaluation system produces side effects. (Dahler-Larsen, 2011). Is it possible to assume that the use of those tools has led to a fatigue of the evaluation itself. It is essential to improve the evaluation process itself and to make the process more ethical.

THE CAPES EVALUATION MODEL

The CAPES evaluation system (Higher Education Personnel Training Coordination) has wide use in Brazil. This model, created between 1976-1977 to record Brazilian graduate activities memory, is nowadays used to evaluate graduate courses.

Although it was created to evaluate graduate courses, both for the accreditation of new courses and for the re-accreditation of existing ones, the superior education institutions have been using it to evaluate individual researchers. This model is criticized since it is based on criteria developed to evaluate hard sciences. Therefore, the CAPES evaluation system is, indeed, a broad regulatory mark for the structure of the academic life as a whole. The analysis

of the CAPES model could clarify the limits of its use, as this brief presentation shows. The items are: a) Program proposal, b) Faculty, c) Students, PhD theses and Master dissertations d) Intellectual production, e) Social insertion, and f) International insertion.

- a) Program proposal: this assesses the alignment between the course and its research lines, as well the infrastructure offered to faculty and students. It is an obligatory item that, in theory, contextualizes the assessment as a whole.

- b) Faculty: this assesses the composition, the educational diversity and experience of the Faculty. This item requires quantitative and qualitative indicators. But, in essence, this item is evaluated based on quantitative parameters.

- c) Students, PHD theses and Master dissertations: this assesses the input and output stream of students, as well the research products developed by the students. It is a problematic item because it does not evaluate the quality of theses and dissertations in terms of liable indicators.

- d) Intellectual production: this assesses the number of publications: articles, books, research reports, technical and technological products, and artistic production. It is challenging to establish indicators for such different areas as Exact & Earth Sciences; Agronomical, Health and Biological Sciences; Humanities and Social Sciences, and Arts. The dominant criteria is articles presented in mainstream journals, ignoring to a large extent the culture of publication of different areas of knowledge.

- e) Social insertion: this assesses the impacts of graduate courses related to social demands. There are no clear indicators to measure and evaluate this item. In fact, a better definition of social insertion is needed.

- f) International insertion: this is also a controversial evaluation parameter. It is certainly a challenging work to measure in quantitative and qualitative approaches faculty and student mobility, international publications, international reputation, among others. On the other side, the international rankings of universities are increasingly used, but it is necessary to have in mind the objectives and the methods to elaborate these rankings to use them adequately (Vogel & Kobashi, 2015).

In a survey developed between 2013 and 2014, it was found out that 70% of the scholar community criticisms are related to the evaluation of Intellectual Production (Vogel, 2015). To obtain fair assessment, it is necessary to build appropriate criteria applicable to different scientific areas, and expressed in qualitative indicators. In fact, many aspects of evaluation of the six above-mentioned items are being relegated to the background in light of the difficulties of creating reliable indicators.

FINAL CONSIDERATIONS

The use of bibliometric indicators to evaluate the scientific activity booming in Brazil is generally based on models designed by international institutions. A change in the evaluation culture is fundamental. In this sense, the Leiden Manifesto (Hicks, Wouters, Waltman, de Rijcke & Rafols, 2015) offers options that could effectively strengthen Brazilian scientific research. According to the Manifesto, the evaluation, to be relevant, should develop inclusive indicators, able to protect socially relevant research, to promote the transparency of the evaluation processes, to recognize and respect the research, publication and citation practices of different knowledge fields, and to recognize the systemic effects of the indicators (Hicks et al, 2015). Finally, it is needed to evaluate the evaluation, in order to adapt the models to the ongoing changes happening in society. In the Brazilian case, it means depending less on the dominant criteria, largely based on North-South economic political relations. It is desirable to

give priority to South-South relations or, more specifically, to understand scientific research as an activity that requires evaluation focusing on social relevance and equity.

REFERENCES

- Bourdieu, P. (1983). O campo científico. In P. Ortiz, R. (Ed.), *Pierre Bourdieu: sociologia* (pp.122-155). São Paulo: Ática.
- Bourdieu, P. (2004). Os usos sociais da ciência: por uma sociologia clínica do campo científico. São Paulo: Unesp.
- Dahler-Larsen, P. (2011). *The evaluation society*. Standford, CA: Standford Business Book.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S. & Rafols, I. (2015). Bibliometrics: the Leiden Manifesto for research metrics. *Nature*, 520, 429-431. Retrieved from <http://www.nature.com/news/bibliometrics-the-leiden-manifesto-for-research-metrics-1.17351>
- Vogel, M. J. M. (2015). *Brazilian graduate evaluation: analysis of items used by CAPES and the criticism of the academic community*. (PhD thesis), Universidade de Sao Paulo, Sao Paulo.
- Vogel, M. J. M. & Kobashi, N.Y. (2015) Graduate evaluation in Brazil: its criteria. In *Proceedings of the 16th Encontro Nacional da ANCIB, (ENANCIB'2015)*, João Pessoa, PB. Retrieved from <http://www.ufpb.br/evento/lti/ocs/index.php/enancib2015/enancib2015/paper/viewFile/3124/1150>



Performance Based Funding and Researchers' Strategies for Grant Applications

David Johann* and Jörg Neufeld*

*johann@dzhw.eu

German Centre for Higher Education Research and Science Studies (DZHW), Department 2, Schützenstraße 6a, 10119 Berlin (Germany)

**neufeld@dzhw.eu

German Centre for Higher Education Research and Science Studies (DZHW), Department 2, Schützenstraße 6a, 10119 Berlin (Germany)

Abstract

This paper aims at investigating research grant application strategies. In particular, it explores (a) the motives underlying researchers' decisions to select a funding institution/programme and (b) the factors influencing the affiliation with specific types of applicants, especially the role of PBF in scholar's research grant application strategies. Drawing on unique data by the DZHW Scientists Survey 2016 it can be demonstrated that (a) the types of applicants are not equally distributed across the different disciplines and (b) the PBF setting does not play a statistically significant role when it comes to the affiliation with the different applicant types.

Aim/Research questions

Performance (formula) Based Funding (PBF) systems are intentionally designed to enhance the productivity, efficiency, and quality of research activities, usually with regard to publications and external funding. Proponents of PBF emphasize that they indeed have the intended positive effects on researchers' productivity/efficiency or at least on greater transparency regarding researchers' performances.

However, critics argue that the market-orientation of publicly funded knowledge production undermines the intrinsically motivated quest for knowledge and reputation. Further, it is stated that PBF systems cause a concentration on mainstream or popular research topics and methods.

Against this background, the question, whether and to what extent researchers' strategies for grant application are affected by the PBF systems that scholars have to face, must be raised. We contribute to this strand of research by systematically investigating whether PBF has an effect on researchers' application strategies.

Data

The analyses are based on the Scientists Survey 2016 conducted by Department 2 of the German Centre of Higher Education Research and Science Studies (DZHW) from March to May 2016. The population is all research staff at German universities (incl. technical/theological universities, universities of education and art/music/medical schools). The sampling followed a two stage cluster design with a 40% proportionate stratified random sample at the first stage (59 out of 152 universities). At the second stage e-mail addresses of the research staff have been compiled from the respective web sites of the selected

universities. The sampling frame encompassed 74,317 persons of which 55,694 have been invited to participate in the survey. A total of 4,844 questionnaires were completed (response rate = 10%).

The comprehensive questionnaire covered the topics third party funding behaviour, general research conditions, and scientific misconduct.

Application strategies, types of applicants, and differences between disciplines

Strategies for research grant application may vary across researchers and disciplines. Some researchers may apply for specific funding programmes because these programmes cover a very particular or a wider range of topics. Others may weigh the costs and benefits by considering the duration and the fairness of the evaluation procedures or the required effort of an application. Once again others may have had a good experience with funding organisations in the past which encourages them to re-apply for another grant. Regarding the differences between disciplines, external funding plays a minor role in the Humanities and Social Sciences compared to the Life Sciences and Engineering. The latter, in turn, differ in the relevance of fundamental and applied research. However, even within the scientific disciplines considerable variation can be observed concerning the individual researchers' application activity (Böhmer et al. 2010).

The variables we consider to explore the motives underlying researchers' decisions which funding institutions/programmes to select are presented in Table 1. Using Factor Analysis (PFA), we identified three underlying dimensions, which can be described as motivations based on (1) the content of the funding institution/programme, (2) past experiences (retrospective evaluation), and (3) cost-effectiveness (see Table 1). For each dimension we calculated the related factor scores.

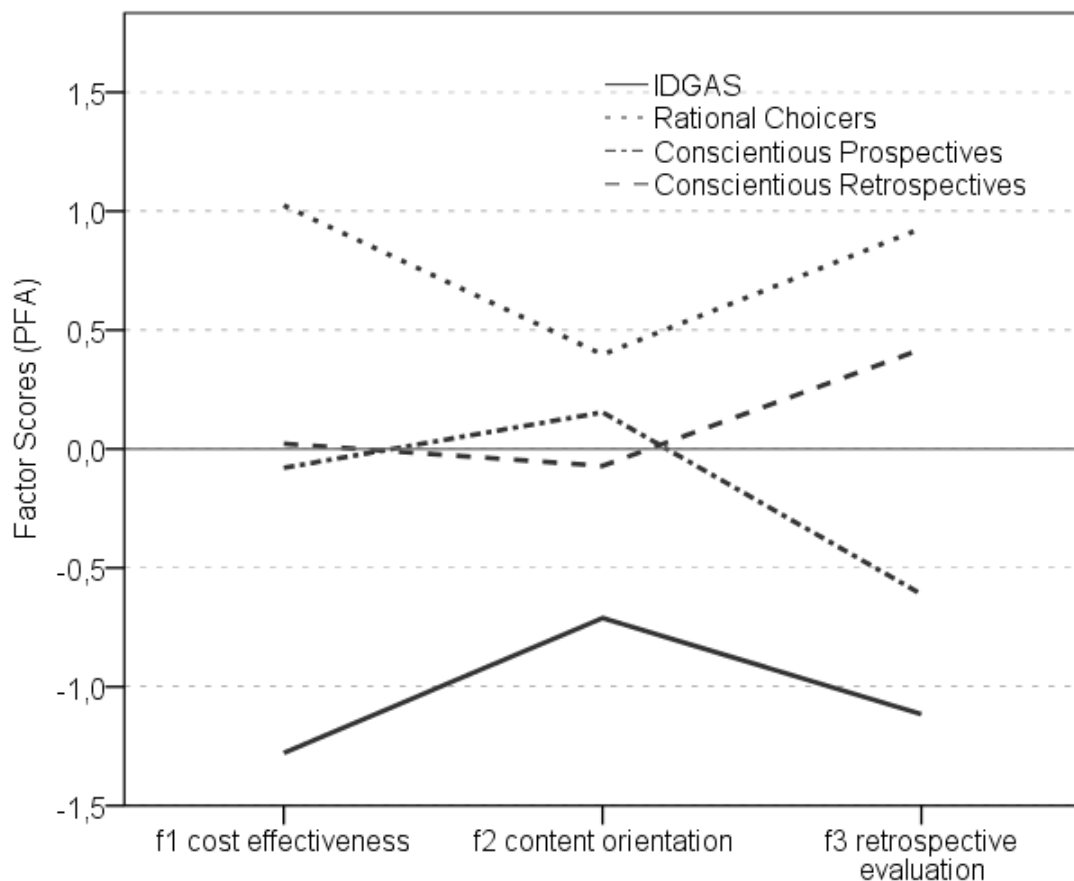
Table 1: Indicators Measuring Different Dimensions of Research Grant Application Strategies

Wording Item	Dimension	Label
The funding programme covers a wide range of topics	1	Content of funding programme
The funding source focuses on fundamental research	1	Content of funding programme
Continuation of previous funding already applied for and granted by the same funding source (follow-up application)	2	Retrospective evaluation
Good experiences in the past	2	Retrospective evaluation
Effort involved in the application	3	Cost-effectiveness considerations
Duration of the evaluation procedure	3	Cost-effectiveness considerations
Fairness of the evaluation procedure	3	Cost-effectiveness considerations
Estimated likelihood of approval	3	Cost-effectiveness considerations

Question wording: *“How important were the following criteria for the decision to submit your most recently granted third-party funding application to the chosen funding source?”*

In order to identify specific types of applicants, we run latent profile analyses (LPA) on the basis of the factor scores (Muthén and Muthén 1998-2010). Even though the BIC and AIC criteria indicate slightly less desirable values in comparison to a typology with five classes, we opt for the four-class solution. The reason for this is that the latent class probabilities (see Table A1) better allocate the respondents to different types of applicants (Geiser 2006). The resulting types of applicants can be described as: (1) *IDGAS¹ Applicants*, i.e., applicants who rely on none of the criteria we consider, (2) *Rational Choicers*, i.e., applicants who primarily consider cost-effectiveness, (3) *Conscientious Prospectives*, i.e., applicants who primarily and equally account for the content of the funding programme as well as costs and benefits of the application, and (4) *Conscientious Retrospectives*, i.e., conscientious applicants accounting for all criteria, but predominantly selecting on the basis of their retrospective evaluations (see Figure 1). As expected, the types of applicants are not equally distributed across the different disciplines (see Table 2).

Figure 1: Means of factor scores by Applicant Types based on LCA



¹ *IDGAS* stands for “I don’t give a shit”.

Table 2: Amount of Applicant Types by Discipline (%)

	Soc. Sciences & Humanities	Life Sciences	Nat. Sciences	Engineering	
Class 1: IDGAS Applicants	19.2% (113)	12.8% (71)	10.0% (40)	11.3% (49)	13.8% (273)
Class 2: Rational Choicers	12.1% (71)	20.9% (116)	17.8% (71)	23.0% (100)	18.1% (358)
Class 3: Conscient Prospectives	34.3% (202)	29.1% (162)	31.6% (126)	23.5% (102)	29.9% (592)
Class 4: Conscient Retrospectives	34.5% (203)	37.2% (207)	40.6% (162)	42.2% (183)	38.2% (755)
Total	100.0% (589)	100.0% (556)	100.0% (399)	100.0% (434)	100.0% (1978)

Note: Number of observations in parentheses. Italic and bold scores mean that the amount is above average

Factors influencing the affiliation with specific types of applicants

Now, we turn to factors influencing the affiliation with specific types of applicants, in particular the role of PBF in scholar's research grant application strategies.

In order to test whether PBF affects researchers' strategies for grant application we calculated multinomial logit regression models. The PBF setting is measured by a sum index indicating how many of the following performance criteria are crucial for the amount of basic research funding the researchers have access to² (1) publications, (2) citation numbers/journal impact factors, (3) third-party funding acquisition (TPF), and (4) results of evaluations.

We include anomie, the perceived general performance pressure, the motivation to apply for funding, and attitudes towards competition in research as control variables. Moreover, we control for gender and age.

The results in Table 3 display average marginal effects of the factors explaining the affiliation with the different types of applicants. It shows that apart from the Life Sciences, the PBF setting does not play a statistically significant role when it comes to the affiliation with the different applicant types. In the Life Sciences the PBF setting increases the probability of belonging to the *Rational Choicers* by four percentage points. However, researchers in the Life Sciences also tend to be *Rational Choicers* when they indicate more pragmatic motives to acquire TPF (see Table B1). The same motives decrease their probability to belong to the *Retrospectives* by six percent.

A further relevant factor is a pro-competitive *attitude*. In the Engineering Sciences as well as in the Social Sciences and Humanities such an attitude decreases the probability of belonging to the *IDGAS applicants* by 4 and 6.5 percent, respectively. It is interesting that the same *attitude* increases the probability of belonging to the *Rational Choicers* in the Engineering Sciences (+7.9 percent) while simultaneously reduces it in the Life Sciences (-5.8 percent).

² Question Wording: "Is some of your basic research funding dependent on the following performance criteria?"

Discussion & Outlook

The analysis has shown that the individual PBF setting seems to have no considerable effect on the criteria researchers adopt when choosing funding organizations in most of scientific fields. Instead, the general necessity of acquiring TPF for pursuing research as well as a pro-competitive attitude indicates statistically significant effects, even though they are rather small.

The questionnaire of the DZHW Scientist Survey 2016 covers many other aspects of scientists' research conditions, personality traits (e.g., the willingness to take risks) as well as their intrinsic and extrinsic motivations for work. However, considering these was beyond the scope of this paper. Yet, these factors may mediate/constitute the perception of performance pressures and are thus promising pathways to better understand grant application behaviour.

Table 3: Effects on the Affiliation with Applicant Types

	Soc. Sciences & Humanities	Life Sciences	Natural Sciences	Engineering Sciences
Class 1: IDGAS Applicants				
PBF Setting	-0,002	-0,012	-0,004	-0,014
Anomie	-0,007	-0,005	0,005	-0,003
Pressure	0,005	-0,009	0,003	0,001
Motivation	-0,017	0,005	-0,001	0,008
Attitudes	-0,039**	-0,008	-0,001	-0,064*
Gender (1=female)	0,004	-0,040	0,020	0,154
Age 2	-0,144	-0,040	-0,024	-0,067
Age 3	-0,138	-0,062	0,056	-0,043
Class 2: Rational Choicers				
PBF Setting	0,011	0,043+	-0,033	0,072
Anomie	-0,007	0,000	-0,005	0,020
Pressure	-0,005	-0,004	-0,003	0,041
Motivation	0,019+	0,069*	0,014	0,004
Attitudes	0,013	-0,058**	0,035	0,079**
Gender (1=male)	0,029	0,100	0,035	0,060
Age 2	0,092	0,190+	0,223**	-0,047
Age 3	0,064	0,214*	0,202***	0,070
Class 3: Conscientious				
Prospectives				
PBF Setting	0,022	-0,027	0,001	-0,026
Anomie	0,019	0,012	-0,013	-0,041
Pressure	-0,014	-0,005	-0,011	-0,041+
Motivation	-0,009	-0,014	-0,025	-0,022
Attitudes	-0,002	0,038+	0,027	-0,027
Gender (1=female)	0,045	0,003	-0,152	-0,184+
Age 2	-0,038	-0,234	-0,131	0,069
Age 3	0,001	-0,246	-0,157	-0,178
Retrospectives				
PBF Setting	-0,031	-0,005	0,036	-0,032
Anomie	-0,004	-0,008	0,013	-0,046*
Pressure	0,015	0,018	0,010	0,000
Motivation	0,006	-0,059*	0,012	0,010
Attitudes	0,028	0,029	-0,061*	0,011
Gender (1=female)	-0,078	-0,063	0,097	-0,029
Age 2	0,090	0,083	-0,075	0,046
Age 3	0,073	0,095	-0,102	0,150
N	159	122	105	72

Note: Coefficients are average marginal effects based on multinomial logit models.

***p<0.001, **p<0.01, *p<0.05, +p<0.10

References

- Böhmer, S., Neufeld, J., Hinze, S., Klode, C., & Hornbostel, S. (2011). *Wissenschaftler-Befragung 2010: Forschungsbedingungen von Professorinnen und Professoren an deutschen Universitäten*. iFQ Working Paper No. 8.
- Geiser, C. (2006). *Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung*. Wiesbaden: VS Verlag.
- Grimpe, C. (2012). Extramural Research Grants and Scientists' Funding Strategies: Beggars Can't Be Choosers? *Research Policy*, 41, 1448-1460.
- Huber, N., Wegner, A. & Neufeld, J. (2015). *MERCI (Monitoring European Research Council's Implementation of Excellence): Evaluation Report on the Impact of the ERC Starting Grant Programme*. iFQ Working Paper No.16.
- Muthén, L. K. & Muthén, B. O. (1998-2010). *Mplus User's Guide. Sixth Edition*. Los Angeles, CA: Muthén & Muthén.

Appendix

Table A1: Average Latent Class Probabilities for Most Likely Latent Class Membership (Row) by Latent Class (Column)

	Class 1	Class 2	Class 3	Class 4
Class 1	0.864	0.000	0.128	0.007
Class 2	0.000	0.790	0.005	0.205
Class 3	0.085	0.008	0.770	0.137
Class 4	0.003	0.144	0.118	0.735



Impact of research evaluation modes of public research funding on the development of research fields and groups in Estonia¹

Kaija Valdmaa* and Piret Tõnurist**

*kaija.valdmaa@gmail.com

Ragnar Nurkse School of Innovation and Governance, Tallinn University of Technology, Ehitajate tee 5, Tallinn 19086 (Estonia)

**piret.tonurist@gmail.com

Ragnar Nurkse School of Innovation and Governance, Tallinn University of Technology, Ehitajate tee 5, Tallinn 19086 (Estonia)

ABSTRACT

The aim of this paper is to analyse how evaluation modes in public funding of research influence the dynamics of public research. The broader theoretical contribution is to develop a model that shows the interplay between different dimensions of research evaluation (e.g. evaluating agency, aim, motivation, time, criteria, tools and impact of evaluation modes) and their influence on the direction of research. Linked with the ancient discussion about scientific excellence versus socio-economic relevance of research there is a contradiction between the ex ante and ex post evaluation modes currently in use in which the long-term salience of research proposed up front is juxtaposed with accountability and legitimacy concerns. Preliminary results of the analysis show that allocating public funding to all research fields under same evaluation criteria produces adverse effects. Main problems are related to reputation, signalling and resource concentration effects that pose a general threat to the sustainability and continuity of a heterogeneous science landscape and contribute to a self-enforcing research system.

AIM, THEORETICAL CONTRIBUTION AND CONSIDERATIONS

The aim of this work in progress is to analyse how evaluation modes in public funding of research influence the dynamics of public research in different research fields and groups (ICT, Biotech, Energy and Environmental Technologies). The main theoretical contribution of this article is to identify relevant dimensions of the main evaluation modes (see Table 1) and develop a model that shows how certain dimensions of research evaluation determine others – what is the interplay between these dimensions and what are the potential effects of evaluation modes (and their dimensions) on research dynamics (see Table 2). Linked with the ancient discussion about scientific excellence versus socio-economic relevance of research there is a contradiction between the ex-ante and ex-post evaluation modes currently in use in which the long-term salience of research proposed up front is juxtaposed with accountability and legitimacy concerns.

¹ This work was supported by Estonian Science Fund grant ETF9395.

Table 1. Dimensions of research evaluation.

Function & aim	Level	Time	Criteria/Scope (of impact)	Method/tool	Indicators
*Summative function	*System	*Ex ante	*Internally focused/Scientific excellence	*Peer review	*In general output measures (volume, quality, impact and utility)
*Formative function	*University	*Ex post		*Bibliometrics	
	*Department			*Co-word analysis	
	*Research group	*On-going	*Wider social and economic context/Socio-economic relevance	*Patent analysis	* Scientific activity, scientific production and scientific progress
*Periodical evaluations	*Project			*Case studies	
*Allocation of public funds				*User surveys	*Policy variables, resource variables, structure variables, process variables, output variables, outcome variables and impact variables
				*Cost/Benefit Analysis etc.	

Source: Authors, based on Geuna and Martin (2003), Arnold and Balázs (1998), Martin (2003), Kogan (1989), Massy (1996), Suter (1997, in OECD, 1997), Cunningham and Gök, (2012), Gök and Edler (2012).

Table 2. Preliminary vision of an interrelated system for analysing research evaluation. Estonian targeted financing system highlighted in grey.

Evaluating community	Function & aim	Motivation	Level of evaluation	Time of evaluation	Criteria	Indicators	Method/tool	Expected effects
						*In general output measures (volume, quality, impact and utility)		*In general scientists change their activities in accordance to the indicators used in evaluation
*Academics	<ul style="list-style-type: none"> *Summative function *Formative function *Periodical evaluations *To allocate public funds 	<ul style="list-style-type: none"> *In general rent-seeking versus looking good in international statistics *Legitimation/Objectification *Relevance *Accountability 	<ul style="list-style-type: none"> *System *University *Department *Research group *Project 	<ul style="list-style-type: none"> *Ex ante *Ex post *Ongoing 	<ul style="list-style-type: none"> *Internally focused/Scientific excellence 	<ul style="list-style-type: none"> *Scientific activity, scientific production and scientific progress 	<ul style="list-style-type: none"> *Peer review *Case studies *Bibliometrics *User surveys *Co-word analysis *Patent analysis 	<ul style="list-style-type: none"> *Skewing effect towards scientific excellence, high technology and natural sciences based research *Weakens the basic science capabilities of applied research groups *"Winner takes it all" reward systems that generate signalling and reputation effects of research grants *Resource concentration effects *Self-enforcing system
*Politicians *?					<ul style="list-style-type: none"> *Wider social and economic context/ Socio-economic relevance 	<ul style="list-style-type: none"> *Policy variables, resource variables, structure variables, process variables, output variables, outcome variables and impact variables 	<ul style="list-style-type: none"> *Case studies *User surveys *Peer review *Cost/Benefit Analysis etc. *Patent analysis 	<ul style="list-style-type: none"> *Pressures for universities to find alternative sources of funding generates substitution and complementary effects *Research could shift closer to the market and further away from creating new knowledge - threat to the sustainability of basic research VS an opportunity for the entrepreneurial university' to become the 'engine' of the knowledge economy

Source: Authors, based on Geuna and Martin (2003), Arnold and Balázs (1998), Martin (2003), Kogan (1989), Massy (1996), Suter (1997, in OECD, 1997), Cunningham and Gök, (2012), Gök and Edler (2012).



FOCUS AND RESEARCH METHOD

We are looking at the effects of the prevalent evaluation mode of research – the excellence-based model. To challenge the applicability of the prevalent mode the case of a small country like Estonia is chosen, which is also an extreme example of an ‘excellence-based’ and also ‘project-based’ science funding system (see e.g. Masso and Ukrainski, 2009; Raudla et al., 2015) and additionally a good case because of the absence of other selection mechanism. Our focus is on the impact of the evaluation mode of the main state funding program for science – Ministry of Education and Research target financing. The main dimensions of this system and its main effects seen by research groups are highlighted in grey in Table 2.

The posed research puzzle is analysed based on the research performance related information (input, output, collaboration) from formal documents available in Estonian Science Information System (ETIS), Estonian Ministry of Education and Research archive and 47 qualitative interviews with research group leaders (35) and representatives of financial (6) and research and development (6) departments from the main public universities in Estonia.

PRELIMINARY FINDINGS

Preliminary results of the analysis show that allocating public funding to all research fields under same evaluation criteria produces adverse effects. Main problems are related to reputation, signalling and resource concentration effects that pose a general threat to the sustainability and continuity of a heterogeneous science landscape and contribute to a self-enforcing research system.

- The Estonian case illustrates well how the competitive excellence-based research funding is focused on selecting the ‘best’ research (themes and teams) in high-technology, predominantly natural sciences based research (e.g. biotechnology related fields). Strong groups with good capabilities and networks in top scientific fields benefit proportionally more from the funding system and experience an expansion through the latter.
- The Estonian case illustrates well how the competitive excellence-based research funding is focused on selecting the ‘best’ research (themes and teams) in high-technology, predominantly natural sciences based research (e.g. biotechnology related fields). Strong groups with good capabilities and networks in top scientific fields benefit proportionally more from the funding system and experience an expansion through the latter.
- Research in more applied fields (energy and environmental technologies) is less sustainable and more dependent on industry contracts and market trends. Without or with little public funding these research groups need to make more contracts with the industry to make up for the lost finances and are therefore less independent in choosing their research fields and securing their basic science input for doing applied sciences – this kind of system will in the long run hollow out the basic science capabilities of these research groups and in the worst scenario these groups and research fields will just slowly die out (an extreme case of energy related research fields in TUT).
- However, there are also groups in more applied fields that have learned to play within the rules of the game and manage to fulfil the requirements of both – public funding

and industry contracts – and follow the scientific criteria and produce needed scientific but also applied outputs.

- The project based focus seems to challenge or deepen the problems of excellence-based funding of science even more. The Estonian case study shows that short-term projects and contracts jeopardise in the long run the continuity and sustainability of the research theme and group.

References

Arnold, E., Balázs, K. (1998). *Methods in The Evaluation of Publicly Funded Basic Research. A Review for OECD*. Technopolis, Brighton, UK.

Cunningham, P. N., & Gök, A. (2012). *Compendium of Evidence on the Effectiveness of Innovation Policy Intervention*. NESTA/MIOIR: London/Manchester.

Geuna, A., Martin, B., R. (2003). University research evaluation and funding: an international comparison. *Minerva*, 41, 277-304.

Gök, A., & Edler, J. (2012). The use of behavioural additionality evaluation in innovation policy making. *Research Evaluation*, 21(4), 306-318.

Kogan, M. (1989). The Evaluation of Higher Education: An Introductory Note, in M. Kogan (ed.), *Evaluating Higher Education*. London: Jessica Kingsley Publishers, 11–25.

Masso, J., & Ukrainski, K. (2009). Competition for public project funding in a small research system: the case of Estonia. *Science and Public Policy*, 36(9), 683-695.

Massy, W., F., (ed.) (1996). *Resource Allocation in Higher Education*. Ann Arbor: University of Michigan Press.

Raudla, R., Karo, E., Valdmaa, K. and Kattel, R. (2015). Implications of project-based funding of research on budgeting and financial management in public universities. *Higher Education*, 70(6), 957-971.

Suter, L., E. (1997). United States: The Experience of the NSF's Education and Human Resources Directorate', in OECD, 1997. *The Evaluation of Scientific Research: Selected Experiences*. Paris: OECD.



The More Funding Sources, the More Citations? The Feasibility Study of Design on “Funding Diversity Indicator”

Carey Ming-Li Chen

carey.mlchen@gmail.com; carey.mlchen@narlabs.org.tw

1. Science & Technology Policy Research and Information Center (STPI), National Applied Research Laboratories (NARLabs), 14F., No.106, Sec. 2, Heping E. Rd., Taipei City 10636 (Taiwan)
2. Graduate Institute of Business Administration, National Taiwan University, No.1, Sec. 4, Roosevelt Rd., Taipei City 106, Taiwan(R.O.C)

This study utilized funding acknowledgement (FA) data to analyze the relationship between research sponsorship and research impact in the selected publications which were partially granted by National Science Council (NSC, but now the name is changed to Ministry of Science and Technology) in Taiwan. Compared to the previous studies, this study combined the FA data with grant information recorded in the database of funding organization to calculate how many granted projects financially supported the publications instead of just identifying whether the specific funding organization was mentioned in FA or not. A total of 40,532 publications were analyzed and the study found that the relationship between the number of NSC granted projects and the citations was positive significantly. This study also segmented publications into three groups based on the number of funding departments to answer whether the more diverse funding sources created higher research impact. The result shown that the publications granted by multiple NSC granted projects and also from multiple funding departments received significant higher citations than those who were granted by single project or by multiple projects but from single funding department. The reason behind might be the researchers who were from different disciplines received the grants from their own funding departments and the collaboration among them might create greater visibility of the publications. Based on the results, this study would like to propose the idea of design on “funding diversity indicator” to help to evaluate or predict the research impact. However, the limitation of utilizing FA data needs to be overcome and it should include more information of granted projects from other foreign funding organizations in order to present more comprehensive analysis to help shape the funding policy in the future.



Accuracy and completeness of funding data in the Web of Science¹

B. Álvarez-Bornstein^{*}, F. Morillo^{**} and M. Bordons^{**}

^{*} belen.alvarez@cchs.csic.es

^{**} fernanda.morillo@cchs.csic.es; maria.bordons@cchs.csic.es

IFS, Centre for Human and Social Sciences (CCHS), Spanish National Research Council (CSIC), Albasanz 26-28, Madrid 28037 (Spain)

ABSTRACT

This study analyses funding information included in scientific publications and recorded in the Funding Agency (FA) and Grant Number (GN) fields of the Web of Science, both extracted automatically from the Funding Text (FX) field. The exhaustiveness and accuracy of the extraction of data from the FX field is analysed in a sample of 1045 papers selected from the Spanish scientific output in 2014. Main mistakes in the extraction process and gaps in the funding data provided by authors and collected in FX are explored. The aim of the paper is to assess the validity of studies based on these WoS subfields; to draw recommendations oriented to improve funding data collection in papers and databases; and, in the long time, contribute to enhance the reliability and usefulness of studies based on funding acknowledgments.

INTRODUCTION

The analysis of funding acknowledgments included in publications is interesting for multiple purposes such as tracking the research output supported by funding bodies and specific grants/research programs; exploring the relationship between research impact and economic support; and identifying the strategic scope of a funding body (Rigby, 2011).

However, the analysis of funding data is a hard task because they are usually included in the acknowledgements of papers, which is a non-structured section with a heterogeneous content: technical, moral or intellectual support can be acknowledged besides the financial one. Accordingly, the automatic extraction of funding data can be difficult and is not always free of errors.

The paragraph containing the funding information as covered in the original paper is recorded in Web of Science in a field named “funding text” (FX). Moreover, the funding organizations and grant numbers are extracted and recorded in two separate subfields: “funding agency” (FA) and “grant number” (GN), which is a valuable improvement to enhance funding acknowledgment studies (Figure 1).

¹ This research was supported by MINECO (grants CSO2014-57826P and FPI BES-2015- 073537)

Figure 1: Example of funding acknowledgements fields in a WoS publication.

FX		
We are grateful for the financial support of this work to the Spanish Ministry of Science and Innovation (project CTQ2011-28967) and to FEDER funds from the European Union		
⇓		
FA-GN entries	FA	GN
FA-GN1	Spanish Ministry of Science and Innovation	CTQ2011-28967
FA-GN2	FEDER funds from the European Union	

Since 2008, when funding acknowledgements were first included in WoS, a number of studies relying on this field has been published (e.g. Costas and Van Leeuwen, 2012; Díaz-Faes & Bordons, 2014) and the analysis of usefulness and caveats of the FX information processed by Thomson Reuters has become particularly relevant (Sirtes, 2013).

OBJECTIVES

Main objectives addressed in this paper are the following:

- to explore the exhaustiveness and accuracy of the extraction of data from the “funding text” field to the FA and GN subfields and to identify main mistakes in this process (database side);
- to detect gaps in the funding data provided by authors and collected in FX (author side);
- to explore to what extent the geographical origin of foreign funding sources is included in the funding text.

Our final objective is to draw recommendations oriented to improve the way in which authors and databases include funding information and enhance the reliability and usefulness of studies based on funding acknowledgments.

METHODOLOGY

A sample of 1045 papers (articles and reviews) is selected from the scientific publications of Spain with funding acknowledgments information in 2014 (N=51,549). A stratified sampling method was used to assure a good representation of eight different broad thematic areas (Agriculture, Biology and Environment; Biomedicine; Physics; Engineering and Technology; Mathematics; Clinical Medicine; Multidisciplinary; and Chemistry).

Manual revision of data collected in FA and GN as compared to those included in the full FX field was conducted to analyze the following aspects:

1. Exhaustiveness of WoS in collecting funding data: % publications in which all the funding agencies and grants mentioned in FX are included in FA and GN. Moreover, where appropriate, the type of lost information was analyzed (FA, GN or both).
2. Variability in the distribution of collected data: in theory, every financial support could be identified through a grant number and the name of the funding agency (FA-GN entry), so one would expect to find so many FA-GN entries in a paper as research grants. However this is not always the case. The following measures were calculated to explore this issue: % FA-GN entries which include a) one funding agency and one grant; b) more than one agency; c) more than one grant; d) no agency; e) no grant. Our concern is to determine to

what extent the average number of funding agencies and/or grants can be derived with precision from the count of FA-GN entries associated to a given publication.

3. Identification of the geographical origin (country/region) of foreign funding. The percentage of FAs in which the international (a country different from Spain) or supranational (e.g. EU) origin of foreign funds can be obtained.

RESULTS

The rate of exhaustiveness of WoS in covering funding information was high, since all the funding agencies and grants included in the FX field were extracted to the FA and GN subfields in 88% of the publications (921).

For the 1045 publications under analysis, a total of 2924 FA-GN entries were collected, that is 2.8 FA-GN entries per publication. Differences between areas were not statistically significant.

An agency was present in almost all the cases (97%), while the grant number was missing in 43% of them, mainly due to the omission of the information in the FX field. Around half of the FA-GN entries (53%) included both, one or more funding agencies and one or more grant numbers.

If different grants from the same agency were acknowledged in a paper, sometimes they were mentioned together in a single FA-GN entry (63%), while other times they were separated in different FA-GN entries (37%). Accordingly, the number of FA-GN entries was not always a precise measure of the total number of grants supporting the research.

On the other hand, co-funded grants were found in 193 publications (18%). In 48 of these papers (25%), the different agencies co-funding the research were collected in a single FA-GN entry while in 145 they were recorded in different entries (75% of papers).

A foreign source of funding was identified in 1071 FA-GN entries (37%). The geographic origin of the source could be clearly derived from data in 51% of the cases (i.e. country names, adjectives of place, etc), being possible to increase substantially this percentage if the analyst had a good knowledge of the main research funding institutions worldwide (e.g. COST, FEDER, ERA, CYTED, or UNESCO –among others- could be identified as supranational; CNRS, CNR or Max Plank could be identified as international).

DISCUSSION AND CONCLUSIONS

Regarding the exhaustiveness of WoS in extracting funding data from the FX field, our study shows that some funding information is lost in 12% of the publications. Concerning the accuracy of WoS in locating funding data in the FA and GN subfields, some variability or lack of consistent criteria in the way that the information is distributed is observed. This reduces the usefulness of specific automatic counts derived from these subfields, such as the number of research grants per paper. The extracting process followed by WoS could be improved with more clear norms concerning the distribution of basic data.

Anyway, both the exhaustiveness and accuracy of the extracting process of funding data is hindered by the incomplete and/or unclear way in which authors include funding information in papers. Detailed norms concerning the way funding support should be acknowledged by

authors in their papers should be established by funding agencies and supported by scientific journals.

REFERENCES

Costas, R., & Leeuwen, T. N.(2012). Approaching the “reward triangle”: General analysis of the presence of funding acknowledgments and “peer interactive communication” in scientific publications. *JASIST*, 63(8): 1647-1661.

Diaz-Faes, A. A., & Bordons, M. (2014). Acknowledgments in scientific publications: Presence in Spanish science and text patterns across disciplines. *JASIST*, 65(9): 1834-1849.

Rigby, J. (2011). Systematic grant and funding body acknowledgement data for publications: new dimensions and new controversies for research policy and evaluation. *Research Evaluation* 20(5): 365-375.

Sirtes, D. (2013). Funding acknowledgements for the German Research Foundation (DFG). The dirty data of the Web of Science database and how to clean it up. In J. Gorraiz, E. Schiebel, C. Gumpenberger, M. Hörlesberger, & H. Moed (Eds.), *Proceedings of the 14th ISSI Conference* (Vol. 1, pp. 784–795). Vienna, Austria: AIT GmbH.



Patent indicators for the Spanish nanotechnology domain¹

Björn Jürgens* & Víctor Herrero-Solana**

* bjurgens@agenciaidea.es

CITPIA Patent Information Centre, Agency of Innovation and Development of Andalusia, Seville, Spain

** victorhs@ugr.es

SCImago-UGR (SEJ036), Universidad de Granada, Granada, Spain

ABSTRACT

The poster describes a patentometric study of Spanish nanotechnology for the years 2004 to 2014. Relevant patent classifications were identified and combined with an established lexical query for nanotechnology (Magrebi et al 2010). As a patent data source the database *Espacenet-Worldwide* from the European Patent Office was used since a previous study from the authors showed that it provided the best data coverage for the purpose of the study (Jürgens, Herrero-Solana 2015). More than 3400 patent records with Spanish authorship were retrieved and after an exhaustive data harmonization process a patentometric analysis was performed using the software tool *Matheo Patent*. For a patent/paper comparison furthermore scientific article data was retrieved from *Scopus*. Subsequently several indicators were generated.

POSTER TEXT

First, Spanish patenting in Nanotechnology was compared to worldwide patenting and publishing. By launching the search query to the total worldwide database and to applicant affiliations of seven important Nano output countries one could see how the Spanish nanotechnology is behaving compared to an international basis and as a result two types of countries could be identified (Figure 1). On the one hand a group comprising the United States, Japan and South Korea where the production of patents is relatively higher than the scientific production. On the other hand a group with the opposite behaviour, which includes especially China and to a lesser extent the UK and Spain. Spain intervenes at 1% of the patents on nanotechnology in the world, but has more than double the representation for scientific papers.

Regarding the Spanish Nanotechnology thematic profile we compared it with worldwide patenting and could identify an above average patenting in the field of nano-medicine and nano-biotechnology (Figure 2). On the contrary we found a deficit in patents related to nano-optics, nano-magnetism and nanotechnologies related to information and communication technologies (ICT). In the field of materials science related to nanocomposites, production is equivalent in relative terms to the rest of the world.

Regarding the co-authorship of inventors from Spain with inventors from other countries most collaboration in nanotechnology patents is done with inventors from the US, followed by

¹ Work supported by the project: “Vigilancia tecnológica de la nanotecnología Española a través de sus patentes”, granted by the Ministry of Economy of Spain (Plan Nacional de I+D+i 2008-2011, code: CSO2012-38801)

Germany, UK and France (Figure 3). By identifying the patent authorities where the applicants file their patents one can see which countries or patent systems were considered of interest for the applicant to protect their invention. As expected from patents with Spanish authorship most patents were filed at the Spanish patent office (ES), but closely followed by filings of PCT applications (WO) at the World Intellectual Property Organization (Figure 4). The third and fourth most important patent filing destination was the US and the European Patent Office (EP). It is interesting to see that China, seems to have overtaken Japan as a more desirable patenting destination for Spanish nanotechnology.

Regarding the top patenting regions & sectors five focal points of nanotechnology patent generation in Spain could be detected with Barcelona and Madrid leading followed by Valencia, Sevilla and La Coruña (Figure 5). By analyzing the patent output according to its applicant's sector affiliation the universities are prevalent, followed by private enterprises, the CSIC and other research centres (Figure 6). The most inventive applicant in terms of patent family counts was the CSIC research centre Instituto de Ciencia de Materiales de Madrid, followed by the Univ Santiago de Compostela and the Univ Sevilla (Figure 7, left). Their inventive strength is an important factor why their correspondent Spanish regions are amongst the top. When considering the number of patent records we do not evaluate the inventiveness but moreover the willingness of the applicant to extend its invention to multiple countries or patent systems in order to extend its protection. It can therefore be seen as a value indicator, showing the effort of the applicant in internationalizing its invention. Regards to the outcome of the ranking we can observe that the CSIC centre Instituto de Ciencia de Materiales de Madrid lowered its rank whereas the Univ Santiago de Compostela and the Univ Sevilla are leading in terms of total published nanotechnology patent documents (Figure 7, right).

In order to measure the effort of internationalization we describe an indicator, which is a ratio between the number of patent registrations (in different offices) and patent families (the invention or innovation itself) and can be used to measure the value of patents. When we analyse the rate of internationalization in Spain, we found that the companies, whose business model are based on the protection of such innovations and therefore are willing to such an effort, present the highest values (Figure 8).

Finally it was of interest to compare the patenting and scientific publishing behaviour in order to identify some kind of correlation (Figure 9 and Figure 10). The top applicants, the universities of Santiago de Compostela and Seville in the right side, followed with some distance from the Univ Politecnica de Valencia. Although the Univ Santiago has the highest patent output, it has a moderate paper output comparing to the other universities (in red). The most productive entity in both, patent families and papers is the Instituto de Ciencia de Materiales de Madrid. The most productive in paper publishing turned out to be the two universities of Barcelona, although the latter have far less patents. Non-university and CSIC research centres which we can point out is the Institució Catalana de Recerca i Estudis Avancats with a relatively high patent and paper output.

POSTER FIGURES

Figure 1: Worldwide nanotech patents vs. papers output

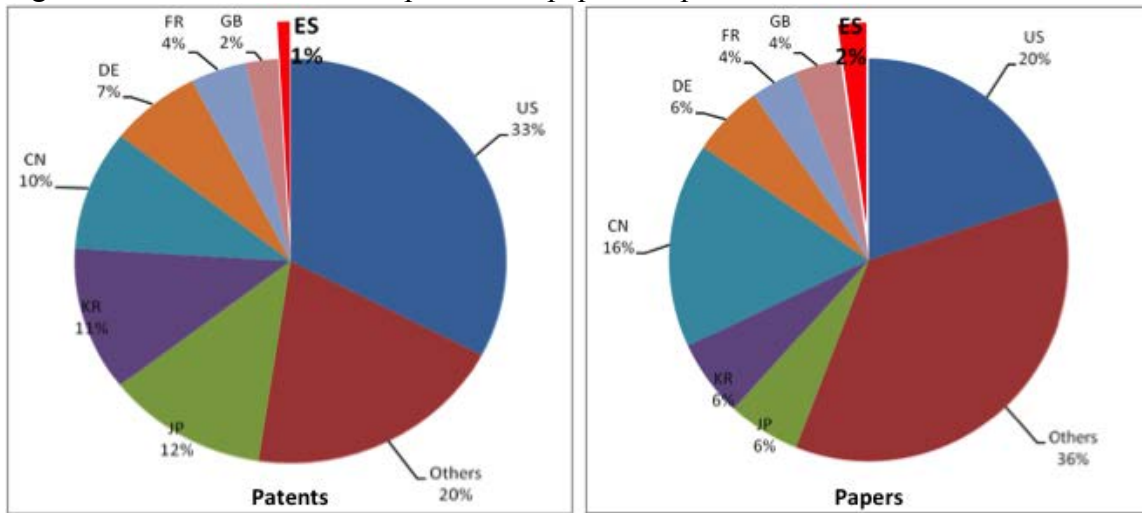


Figure 2: Nanotechnology specialization Spain (left) vs. worldwide patent production (right)

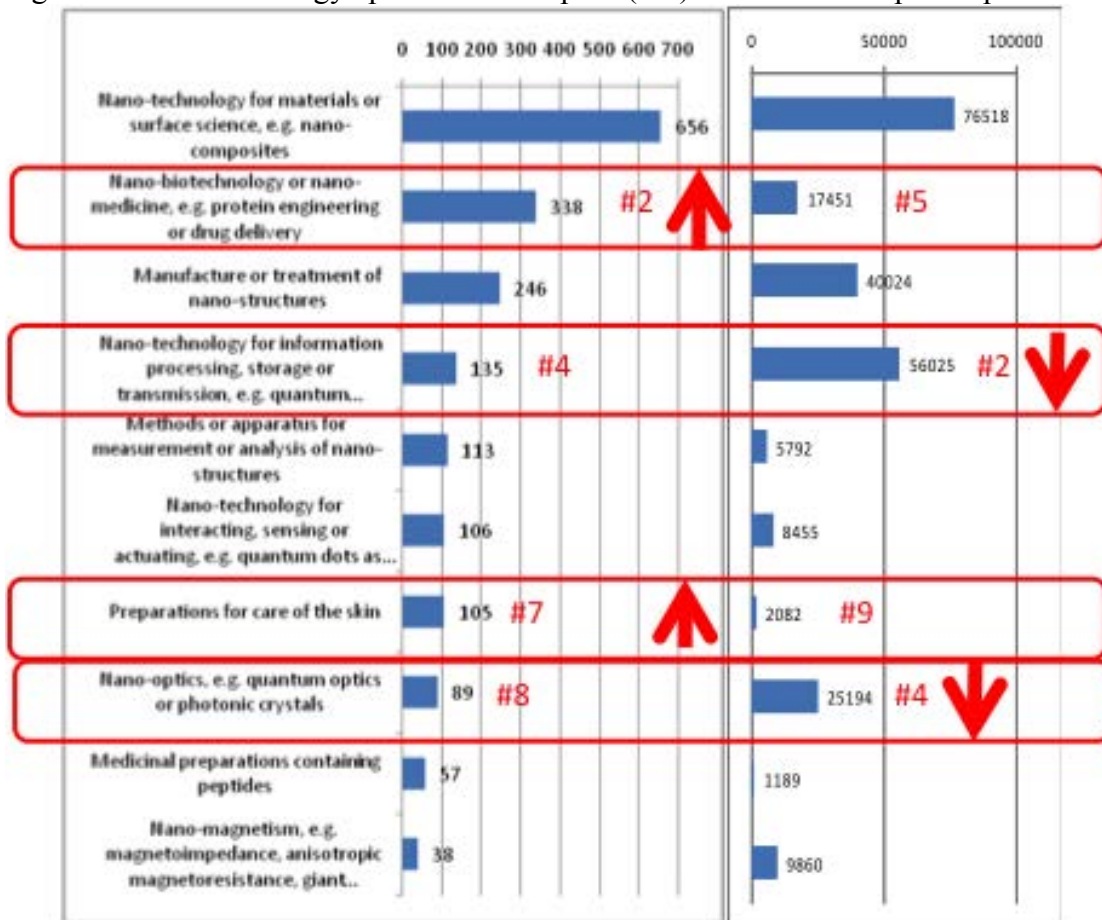


Figure 3: Country of inventors co-occurrence (per family, min. 2 pairs)

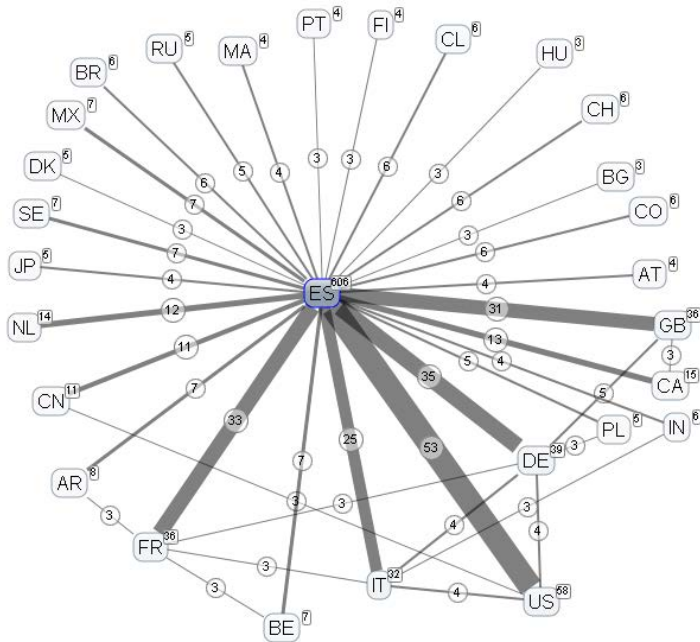


Figure 4: Publications per patent office



Figure 5: Top patenting regions & sectors (per patent family number)



Figure 6: Sector affiliation

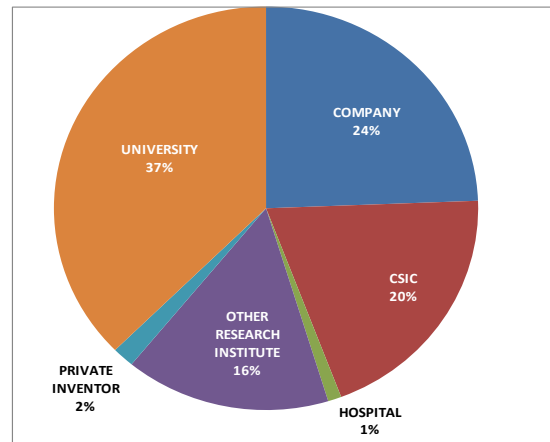


Figure 7: Top 15 Spanish applicants ranked per patent family counts (left) and patent publication records (right)

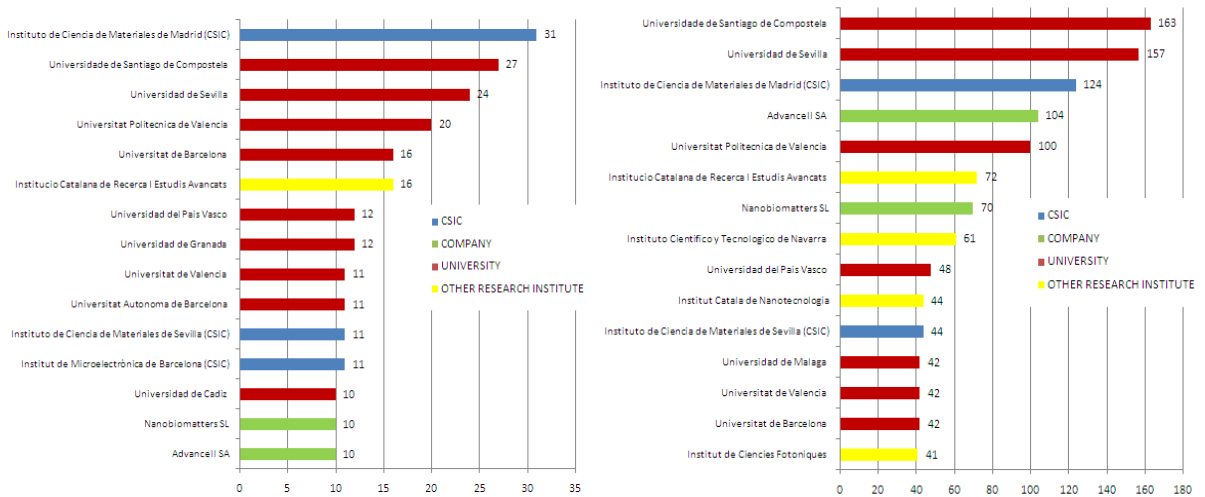


Figure 8: Patent internationalization ratio (IR) of institutions

Applicants (All sectors)	Patent families	Patent records	Patent IR (records/family)
Salvat Lab SA	1	20	20
Grifols SA	2	35	17,5
Interquim SA	1	14	14
Silicalia SL	2	26	13
Dendrico SL	1	13	13
Tcd Pharma SL	1	13	13
Nylstar SA	1	12	12
Hospital De La Santa Creu I Sant Pau	1	11	11
Hospital Universitari Germans Trias I Pujol	1	11	11
Tolsa SA	1	11	11
Avancell SA	10	104	10,4
Biolan Microbiosensores SL	1	10	10
Histocell SL	1	10	10

Figure 9: Article and patent output (per family) of applicants per sector

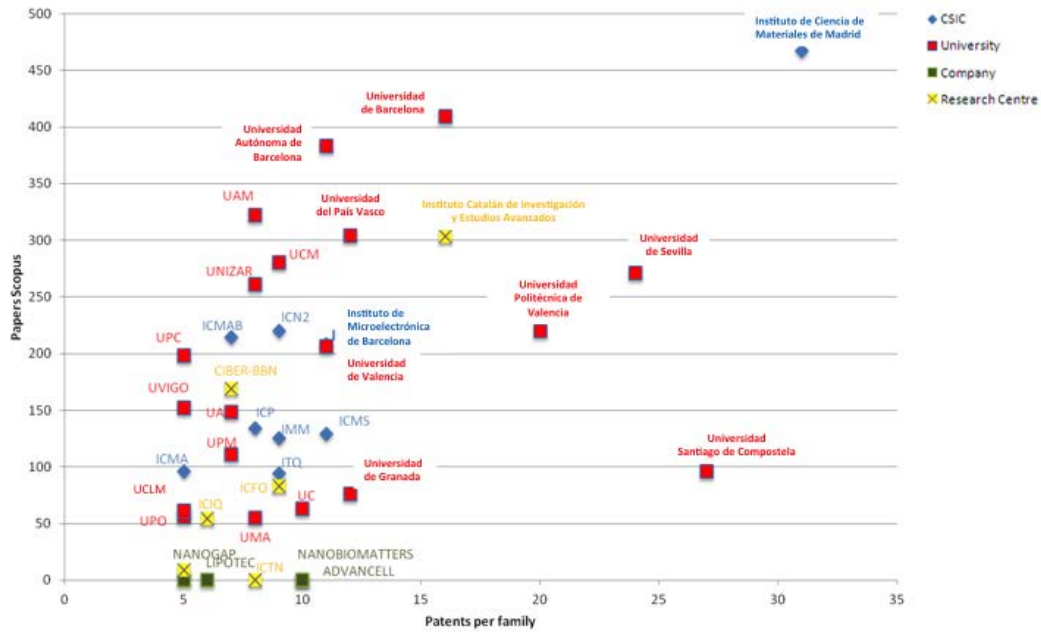
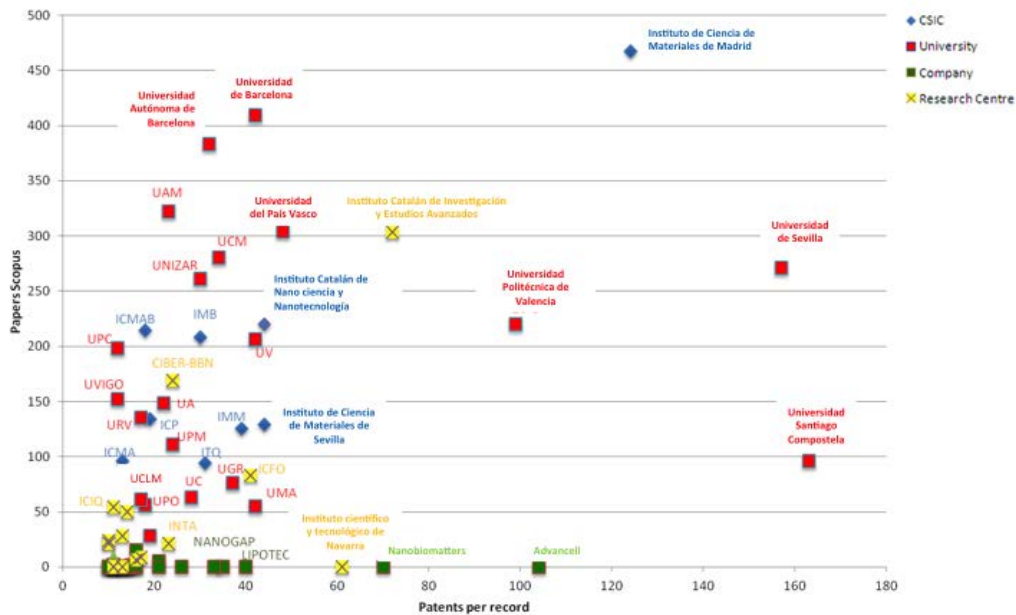


Figure 10: Article and patent output (per records) of applicants per sector



REFERENCES

Maghrebi, M., Abbasi, A., Amiri, S., Monsefi, R., & Harati, A. (2010). A collective and abridged lexical query for delineation of nanotechnology publications. *Scientometrics*, 86(1), 15-25.

Jürgens, B., Herrero-Solana, V. (2015). Espacenet, Patentscope and Depatisnet: A comparison approach. *World Patent Information*. Vol. 42. doi:10.1016/j.wpi.2015.05.004



Best-Practice Benchmarking for Israel: The SNI Scorecard - A Multidimensional Perspective

Shlomo Maital^{*}; Tsipy Buchnik^{*}; Daphne Getz^{*}

^{*} smaital@technion.ac.il; zipi@sni.technion.ac.il; daphne@sni.technion.ac.il
S. Neaman Institute for National Policy Research, Technion city, Haifa, 32000 (Israel)

ABSTRACT

Well-run businesses regularly and systematically benchmark their key business processes against the best practices of other organizations, both within their industry and outside it. Countries, too, are businesses. Countries should also practice best-practice benchmarking, as a fundamental policy tool, by asking two simple questions: What do other nations do better than we? And how can we adapt and adopt what they do, to improve the wellbeing of our citizens? Knowledge of best practices is in general not privileged or secret, is widely available, yet is significantly underused by countries, including Israel, even though the benefits of using it can be striking. In this report, we suggest a simple method for regularly measuring Israel's performance against that of other countries, in 13 dimensions, the SNI Scorecard, as a foundation for shaping policy measures that can significantly improve the wellbeing of the citizenry.

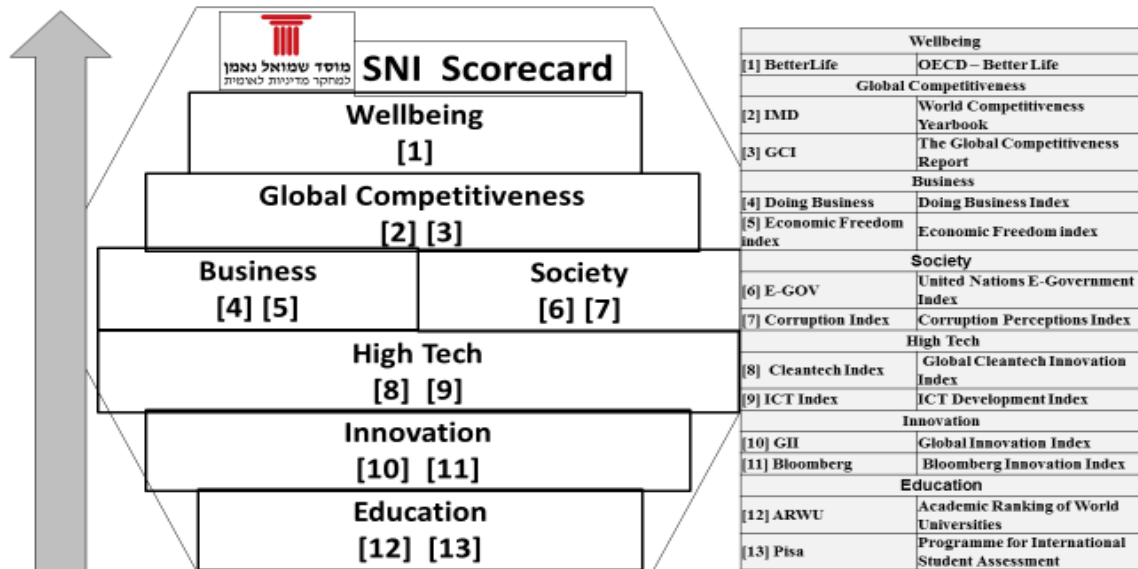
A HIERARCHY OF MEASURES: THE SNI SCORECARD

In this report, we offer a somewhat different and more comprehensive approach to national best-practice benchmarking. We call it the SNI Scorecard. The objective is to systematically benchmark Israel's ranking in the world, compared to other nations, and identify those nations that excel in areas vital to the wellbeing of Israeli citizens. The next step is to ask, what can Israel learn from these nations, and what are their best practices that can be adapted and adopted?

It is self-evident that policy begins with measurement. What you cannot measure, you cannot effectively change and improve. Once, not long ago, it was a difficult task to assemble comparative data on countries' performance. Today, the problem is the opposite – how to sift through mountains of data and choose appropriate measures, then present them clearly, understandably and visually, as a foundation for effective policy debate and formulation.

Figure 1 presents a hierarchy of wellbeing, inspired by psychologist Abraham Maslow's "hierarchy of needs". At the top we place Individual and Family Wellbeing, a multivariate measure that reflects economic and social 'ingredients' of individual and family happiness and welfare. At the foundation, education, which drives innovation, which in turn fuels high-tech industry, in turn a driver of business success and social success, which combine to influence and determine global competitiveness, which ultimately impacts wellbeing. Each of the six 'drivers' of individual and family wellbeing are measured by data from two different databases, in order to provide a full and comprehensive picture.

Figure 1. SNI Scorecard: Israel’s Relative Performance, Benchmarked in 13 Measures

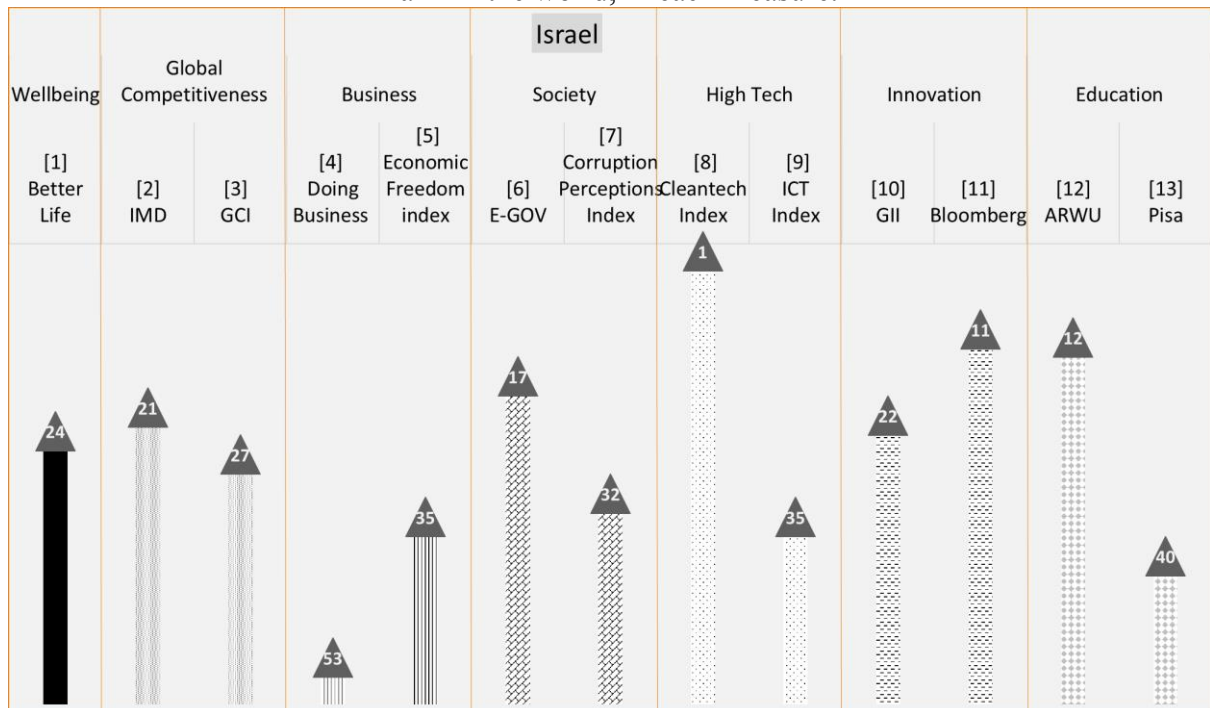


ISRAEL 2016: BEST, GOOD AND WORST

Using the 13 databases, we present Israel’s ranking, in each of the six Scorecard components. Each ranking is calculated by combining a large variety of individual measures, and aggregating the results into a single rank measure. For each of the six components that drive wellbeing, we have used two different databases. For example, for Innovation, we use both the Global Innovation Index, compiled by INSEAD, Cornell and WIPO (World Intellectual Property Office), and the Bloomberg Innovation Index, compiled by Bloomberg Business Week.

The results are shown in Figure 2. This is the SNI Scorecard.

Figure 2. SNI “Hierarchy of Wellbeing” Scorecard, Israel, 2015. The numbers reflect Israel’s rank in the world, in each measure.



This work is licensed under a Creative Commons License: Attribution-NonCommercial-NoDerivatives 4.0 International.

Wellbeing: Israel ranks 24th in the OECD comprehensive measure of wellbeing, out of only 34 member nations of the OECD, representing the main developed countries of the world. This relatively low ranking reflects the exceptionally high variance in the rankings of the components of wellbeing, ranging from first in the world in Cleantech (water and environment), 11th in innovation (Bloomberg) and 12th in university education (ARWU), down to 40th in PISA (high school understanding of reading, math and science), and 53rd in “ease of doing business”.

Suppose Israel’s policy makers want to act effectively and rapidly, to improve the wellbeing of Israel’s citizens. What can they do?

Perhaps, begin by asking, what areas that shape and contribute to wellbeing are relatively weak in Israel. Which countries are relatively strong in these areas? What precisely do they do, that Israel should but does not do?

The benchmarking questions that should be asked on a regular basis by policymakers are:

- **Why** do Australia, Sweden, Switzerland Norway and Denmark lead in overall “wellbeing”? Obviously, Israel would prefer to have as its neighbors, France, Germany and Italy, as Switzerland does, rather than Lebanon and Syria and Gaza. Some elements of Israel’s situation cannot be changed or improved. But others can. What can Israel improve, by learning from nations that rank higher? (see Table 1).
- **Why** are the U.S., Switzerland, Hong Kong and Singapore economies more globally competitive than Israel? What can be learned and improved?
- **Why** does Israel rank low in high-school math and science? What can be improved? How?

Table 1. Best Practice Nations in 13 Score Dimensions: 2015.

	Leading Nation				
	1	2	3	4	5
[1] Better Life	Australia	Sweden	Norway	Switzerland	Denmark
[2] IMD	USA	Hong kong	Singapore	Switzerland	Canada
[3] GCI	Switzerland	Singapore	USA	Germany	Netherlands
[4] Doing Business	Singapore	New Zealand	Denmark	Korea	Hong kong
[5] Economic Freedom index	Hong kong	Singapore	New Zealand	Switzerland	Australia
[6] E-GOV	Korea	Australia	Singapore	France	Netherlands
[7] Corruption Perceptions Index	Denmark	Finland	Sweden	New Zealand	Netherlands
[8] Cleantech Index	Israel	Finland	USA	Sweden	Denmark
[9] ICT Index	Korea	Denmark	Iceland	United Kingdom	Sweden
[10] GII	Switzerland	United Kingdom	Sweden	Netherlands	USA
[11] Bloomberg	Korea	Germany	Sweden	Japan	Switzerland
[12] ARWU	USA	United Kingdom	Switzerland	Germany	France
[13] Pisa	China	Singapore	Hong Kong	Taiwan	Korea

CONCLUSION:

“...the principle of comparing data in the pursuit of performance improvement has become an increasingly normal management activity”.

[Francis & Holloway, 2007, p. 185].

Best practice benchmarking became “an increasingly normal management activity” over a decade ago – but is not yet an everyday ‘normal’ activity among national policy makers in Israel. Yet the potential benefits are very large.

Take, for example, the case of Georgia, a former part of the U.S.S.R. with a heritage of Soviet-era red-tape and bureaucracy. With strong leadership, Georgia systematically benchmarked other nations (e.g. Singapore), that lead in Ease of Doing Business (one of the 13 benchmark measures in the SNI Scorecard), then implemented sweeping and rapid reforms. The result was a dramatic improvement (to 24th in Ease of Doing Business) in Georgia’s economy and in the wellbeing of its citizens.

Comparing Israel – indeed, every country -- with nations that do better should be a regular, systematic and transparent part of policymaking. There is a vast number of accessible databases that make this very easy.

In general, Israel must do better to become a country that learns continually from other countries. To do less is to overlook major benefits that can accrue to the people of Israel, if policymakers would simply look out the window and act on what they see.

REFERENCES

Center for World-Class Universities of Shanghai Jiao Tong University. (2016). *ARWU World University Rankings 2015 Academic Ranking of World Universities*. Retrieved from <http://www.shanghairanking.com/>

Cornell University, INSEAD, and WIPO. (2015). *The Global Innovation Index 2015: Effective Innovation Policies for Development*. (Soumitra Dutta, Bruno Lanvin, Eds.) Retrieved from http://www.codespring.ro/wp-content/uploads/2012/11/GII-2012_Cover.pdf

IMD World Competitiveness Center. (2015). *Competitiveness of Nations and World Economy Ranking*. Retrieved March 8, 2016, from <https://www.imd.org/wcc>

International Telecommunication Union. (2015). *Measuring the information society: The ICT development index 2015*. International Telecommunication Union. Retrieved from <http://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2015/MISR2015-w5.pdf>

Jamrisko, M. and Lu, W. (2016). *These Are the World’s Most Innovative Economies*. Retrieved from <http://www.bloomberg.com/news/articles/2016-01-19/these-are-the-world-s-most-innovative-economies>

OECD. (n.d.). *OECD Better Life Index: Israel*. Retrieved, from <http://www.oecdbetterlifeindex.org/countries/israel/>

- OECD. (2014). *PISA 2012 Results in Focus. Programme for International Student Assessment*. doi:10.1787/9789264208070-en
- Parad, Michele Henningsson, Stefan Currás, T. Youngman, R. (2014). *The Global Cleantech Innovation Index 2014*. Cleantech Group. Retrieved from http://www.cleantech.com/wp-content/uploads/2014/08/Global_Cleantech_Innov_Index_2014.pdf
- Schwab, K., Sala-i-Martin, X., & Brende, B. (Eds.). (2016). *The Global Competitiveness Report 2015-2016* (Vol. 5). World Economic Forum. doi:92-95044-35-5
- The Heritage Foundation. (2016). 2016 Index of Economic Freedom: Israel. Retrieved March 8, 2016, from <http://www.heritage.org/index/country/israel>
- Transparency International. (2015). *Corruption Perceptions Index 2015*. Transparency International. Retrieved from <http://www.transparency.org/cpi2015>
- United Nations Department of Economic and Social Affairs. (2014). *E-Government Survey: e-Government for the Future We Want*. United Nations. United Nations. doi:10.1016/S1369-7021(02)00629-6
- World Bank. (2014AD). *Doing Business 2015: Going Beyond Efficiency*. World Bank. doi:10.1596/978-1-4648-0351-2



Measuring Global Innovation Activities with Article Visiting Geographical Data¹

Xianwen Wang^{*}, Zhichao Fang^{**}, Yang Yang^{**}, Hongyin Wang^{**} and Zhigang Hu^{**}

[*xianwenwang@dlut.edu.cn](mailto:xianwenwang@dlut.edu.cn)

WISE Lab, Dalian University of Technology, Dalian, 116024 (China)

^{**} fzc0225@163.com; yangyang0477@mail.dlut.edu.cn; 361858362@qq.com; huzhigang@dlut.edu.cn

WISE Lab, Dalian University of Technology, Dalian, 116024 (China)

ABSTRACT

In this study, we propose a novel way to analyze the geographical distribution of the global innovation activities. Using 59.64 million article views and downloads to 23.8 thousand scientific papers published by Frontiers, we analyze the worldwide distribution of innovation activities, the knowledge diffusion embodied with different visiting patterns to new or old scientific papers. We find that most innovation activities are concentrated around major metropolitan areas. In most cases, the publishing areas and a few cities are the diffusing source of scientific knowledge, and there exists the diffusing route from the publishing cities to a broader area.

INTRODUCTION

Measuring innovation activities of regions, including the geographical distribution, is an important issue in science policy field. In previous studies, the bibliographic and citation data are aggregated at different scales and used to characterize the geospatial distribution of innovation activities and quantitatively map science clusters in physical space. Frenken et al. (2009) proposed that these kinds of studies could be collected under the general field of ‘spatial scientometrics’. Top continents, countries and cities for scholarly research are identified and visualized by publication ranking algorithm, network-based citation analysis or even journals (Zhang, et al., 2013; Mazlounian, et al., 2013). Centers of excellence around the world are mapped and the distribution patterns and networks of relations are drawn based on data harvested from Scopus or Web of Science through a range of visualization programs (Leydesdorff & Persson, 2010; Bornmann, et al., 2011; Bornmann & Leydesdorff, 2011; Bornmann & Waltman, 2011).

Most innovation activities are concentrated around major metropolitan areas, especially capital regions (Matthiessen, et al., 2010). For example, the Nordic capitals of Hovedstaden, Helsinki-Uusimaa and Stockholm, or the German and Austrian capitals of Berlin and Wien (Kotzeva, et al., 2015). These analyses might help to identify the factors that drive successful research clusters — indicating rising stars and aiding city planners and policy-makers in building profitable centers elsewhere (Richard, 2010).

In these previous studies, bibliographic data and citation data are used to represent the innovation activities to perform the analysis, however, only relying on bibliographic data and

¹ This work was supported by the project of “National Natural Science Foundation of China” (61301227), the project of “Growth Plan of Distinguished Young Scholar in Liaoning Province” (WJQ2014009), and the project of “the Fundamental Research Funds for the Central Universities” (DUT15YQ111). An extension of this study is submitted to *Scholarly Metrics and Analytics* as an inaugural article.

citation data to investigate innovation activities ignores the prepositive procedures. Besides the bibliographic data and citation data, innovation activities could be reflected by many other types of data, accessing to scientific papers could be regarded as an important part of science activities. Few attention has been paid on the behaviours that researchers getting access to scientific papers because of the availability of article usage data. Nowadays, many changes have taken place in scholarly publishing. How many times a scientific paper is used, how users access to the scientific paper and even where the users from? More and more academic services are opening these data to public, like Web of Science, *PLoS*, *Nature*, *Science*, etc. Many journal and publishers disclose the article views/downloads and integrate altmetrics data on the webpage for each article, and some even go further. For example, *Frontiers* disclose the demographics data, which makes it possible for readers to know where the visitors come from.

The access to the large amount of usage dataset brings opportunities to researchers to carry out innovative studies that could not be done by bibliographic records or citation data. Just as Kurtz et al. (2005) concluded, the existence of this information has great implications for the future of information retrieval and bibliometrics. Usage data do constitute a much more comprehensive record of innovation activities than previous bibliographic data (Kurtz & Bollen, 2010). In previous studies, usage data have been used to evaluate researches (Davis, et al., 2008), assess or predict impact (Kaplan, et al., 2000; Brody, et al., 2006; Shuai, et al., 2012; Wang, et al., 2014a), identify usage patterns or rules (Wang, et al., 2014b), detect and trace research trends (Wang, et al., 2013a), and explore user behaviors (Davis & Solla, 2003; Davis & Price, 2006). Combined with the geographical coordinates information, usage data could reflect the diverse user behaviors in different regions (Wang, et al., 2012; Wang, et al., 2013b). As a consequence, usage data of scientific papers provide a new way to reveal innovation activities and knowledge diffusion in the aspect of geographical space.

In the field of scientometrics, using the classical methods, like co-word analysis, co-authorship analysis, citation analysis and co-citation analysis, the diffusion of knowledge could be explored to some extent (Cottrill, et al., 1989; Li, et al., 2007; Neff & Corley, 2009; Naumis & Phillips, 2012; Liu, et al., 2015). In particular, citation relationships between publications are the best illustrations of the knowledge diffusion. But citation data need a long time to accumulate and citation relationships also need a long time to establish due to the citation delays. In addition, bibliographic and citation data only record a little part of the process of knowledge diffusion, traditional scientometrics methods only tell the incomplete story about knowledge diffusion.

In virtue of the visiting data supplied by Frontiers, our research questions are, what is the geographical distribution of visitors to scientific papers? What is the relationship between scientific production and consumption for regions? Is the geographical distribution complying with the distribution of the origins of scientific knowledge? What is the geographical flow of new scientific knowledge?

DATA AND METHODOLOGY

The article usage data is collected from Frontiers directly, which is founded in 2007 and now one of the largest open-access publishers. As of the end of 2015, Frontiers has published 38,000 articles, and receives 101 millions of article views, as shown in Figure 1.

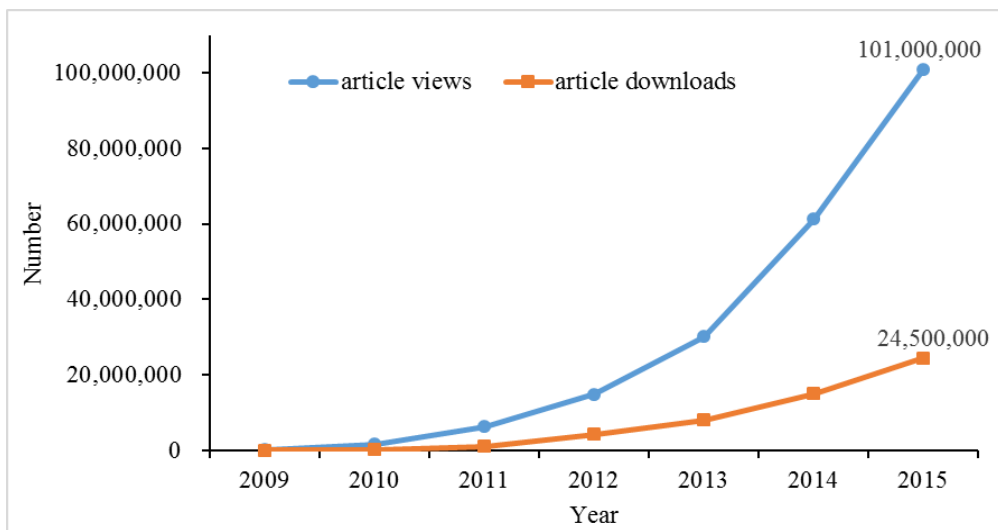


Figure 1 Articles views and downloads of Frontiers articles

In 2008, Frontiers was the first publisher to conceive and provide Impact Metrics to public, known as article-level metrics. Other than the metrics of views, citations and social buzz (through altmetrics.com data), Frontiers also expands its collection of Impact Metrics to the readership demographics for each accepted or published articles. As Figure 2 shows, the geographic summary displays the number of visits where the visitors come from.

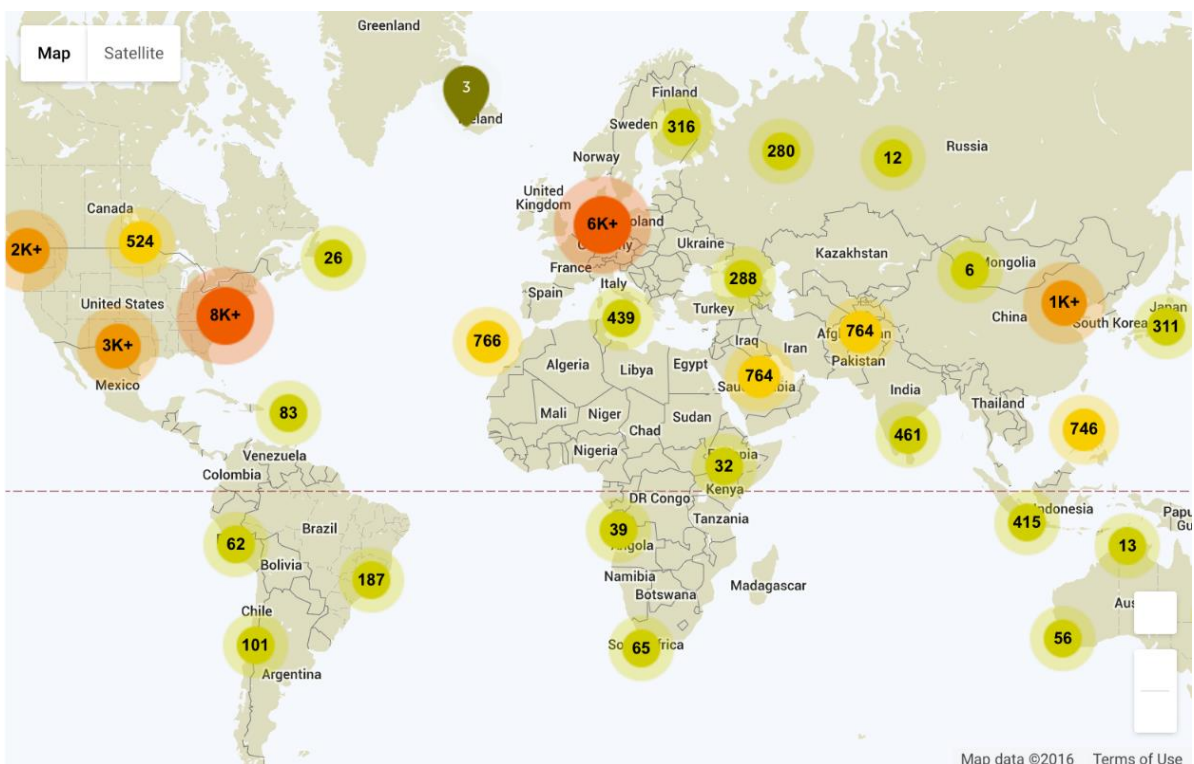


Figure 2 Screenshot of geographic summary of Frontiers

In this study, 16 journals published by Frontiers and indexed in Web of Science are selected as our research objects, as shown in Table 1, 16 journals are ranked by the number of published articles. The records of 23,798 articles published by the 16 journals from 2007 to 2015 are downloaded from Web of Science, while the visiting data of each article is collected directly from the website of Frontiers. 23,798 articles received 59,637,047 article views and downloads. Finally, the publishing data and visiting data are parsed into a special designed SQL database to be stored, processed and analyzed. For the article visiting data harvested from Frontiers, addresses are converted from geographic coordinates (latitude and longitude of the visiting data) with Google Maps Geocoding API. So, the city names and country names are confirmed according to the results returned by the Google geocoder. In our dataset, there are 132,322 unique geographic coordinates need to be converted.

Table 1 dataset used in this study

Rank	Journal	Articles	Visits
1	Frontiers in Psychology	4764	11,613,606
2	Frontiers in Microbiology	3150	6,850,251
3	Frontiers in Human Neuroscience	3055	9,124,859
4	Frontiers in Plant Science	2651	5,947,421
5	Frontiers in Physiology	1768	4,287,535
6	Frontiers in Neuroscience	1398	4,304,982
7	Frontiers in Behavioral Neuroscience	1264	2,896,053
8	Frontiers in Cellular Neuroscience	1262	2,673,111
9	Frontiers in Pharmacology	955	2,448,095
10	Frontiers in Aging Neuroscience	715	1,690,454
11	Frontiers in Computational Neuroscience	701	1,884,851
12	Frontiers in Neural Circuits	599	1,637,839
13	Frontiers in Neuroanatomy	528	1,663,495
14	Frontiers in Cellular and Infection Microbiology	526	1,073,133
15	Frontiers in Molecular Neuroscience	302	998,779
16	Frontiers in Neuroinformatics	160	542,583

RESULTS

Visiting regions

Figure 3 displays visits from 83,195 cities worldwide, the size of nodes represents the number of visits, and nodes are placed on the map according to the latitude and longitude coordinates of the cities. Besides the 20 large nodes with the most visits and labelled with the exact numbers, the dense distribution in the eastern United States, the western Europe, the whole Japan, and even India is rather considerable. Although China has the second most visits in the world, it seems that a few big cities contributed a large proportion of the total visits, while most cities in China remain blank on the map.

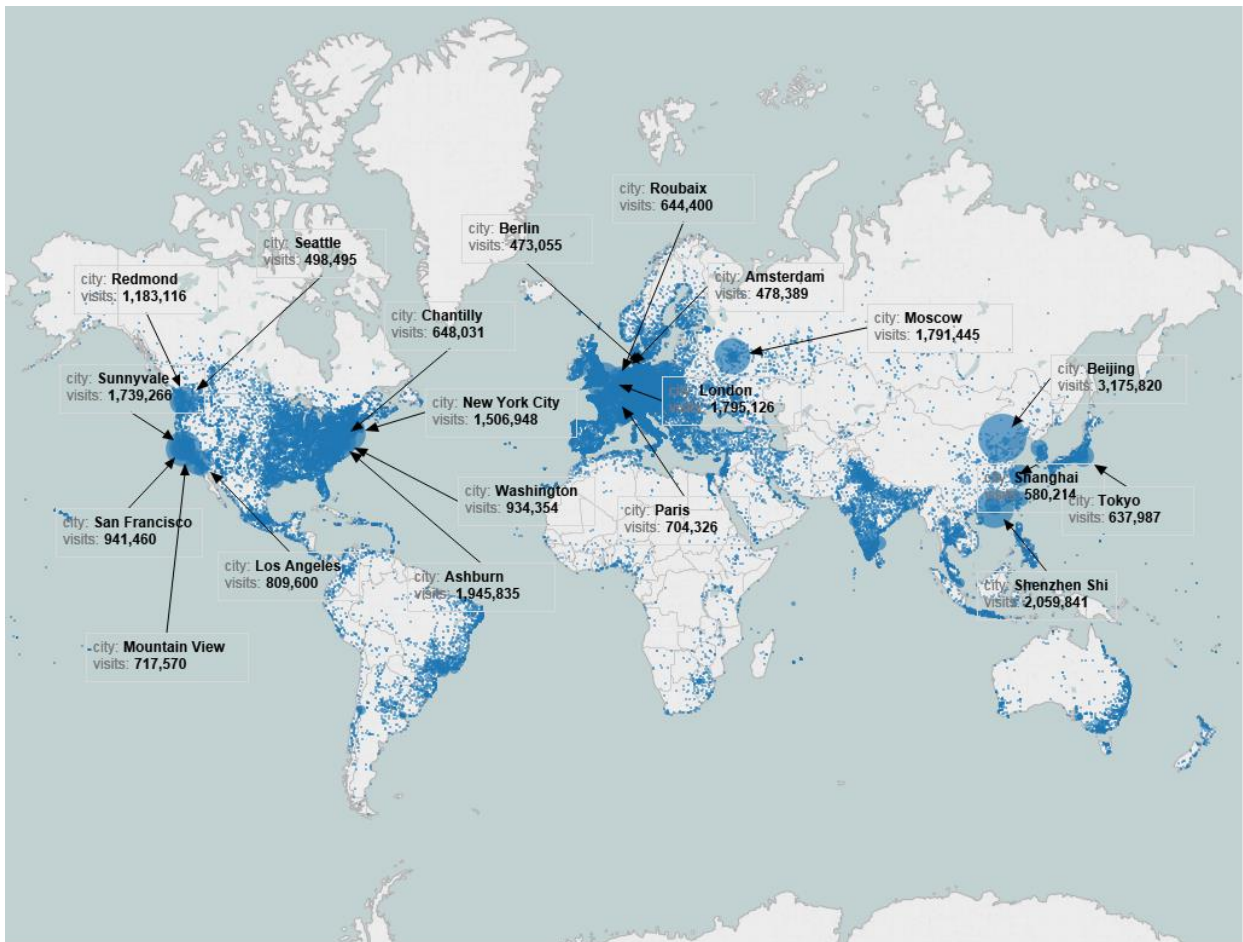


Figure 3 World map of visiting cities

Figure 4 shows the fitting distributions of article visiting data of the global cities. The inset distributions divide the publications into two groups according to the publish date and show the visiting data of cities. All the distributions are fitted well to power law distributions (dashed lines), with the fitted parameter $\alpha = 1.62$. The blue line indicates the overall visiting city distribution, from which we could see, a small part of cities accumulated most article visits, while most cities only visits a few times.

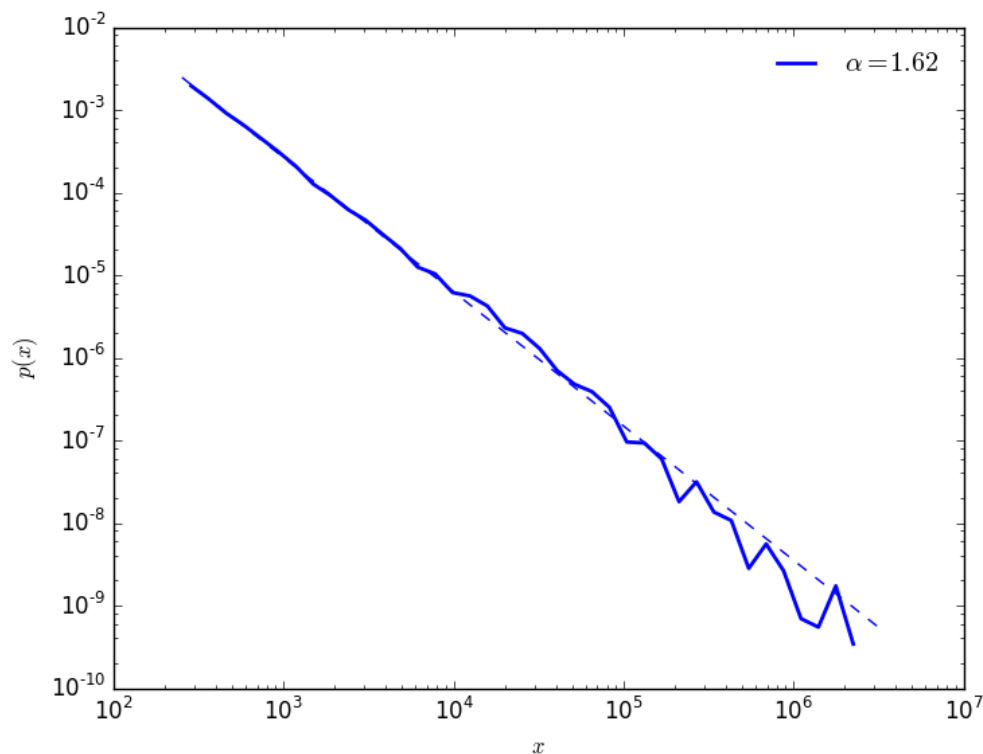


Figure 4 Power law distribution of visiting cities

Visiting geographical distribution to new or old papers

What are the visiting patterns of regions to scientific papers? Are the visiting patterns different between the new papers and old papers? In this section, some newly published papers (121 papers published in late December 2015) and old papers (85 papers published in November – December 2012) are selected to analyze. Because the overwhelming number of visits from USA and China may cause bias to the results, especially for those articles published by the two countries, so articles published by USA and China are excluded here. The results are shown in Figure 5 and Figure 6.

Each bar has all the visiting cities to one particular article, the length of stacked bars represents the number of visits of the cities. The closer to the top, the more the city visits the paper. The publishing city is in red colour, the visiting cities in the same country of the publishing city are in orange colour, and other visiting cities are in grey colour. As Figure 5 shows, most of the 121 newly published papers have the publishing city (the data marker in red colour) in the top. For most articles, the data markers in orange colour are distributed dispersedly from the top to the bottom of the bar. And the area occupied by the red and orange data markers is rather considerable. Besides, although some are not the publishing cities, they are in the top visiting list for many articles. For example, 45 papers have Ashburn from USA, 46 papers have Shenzhen from China and 38 papers have Beijing from China in the top 5 visiting cities.

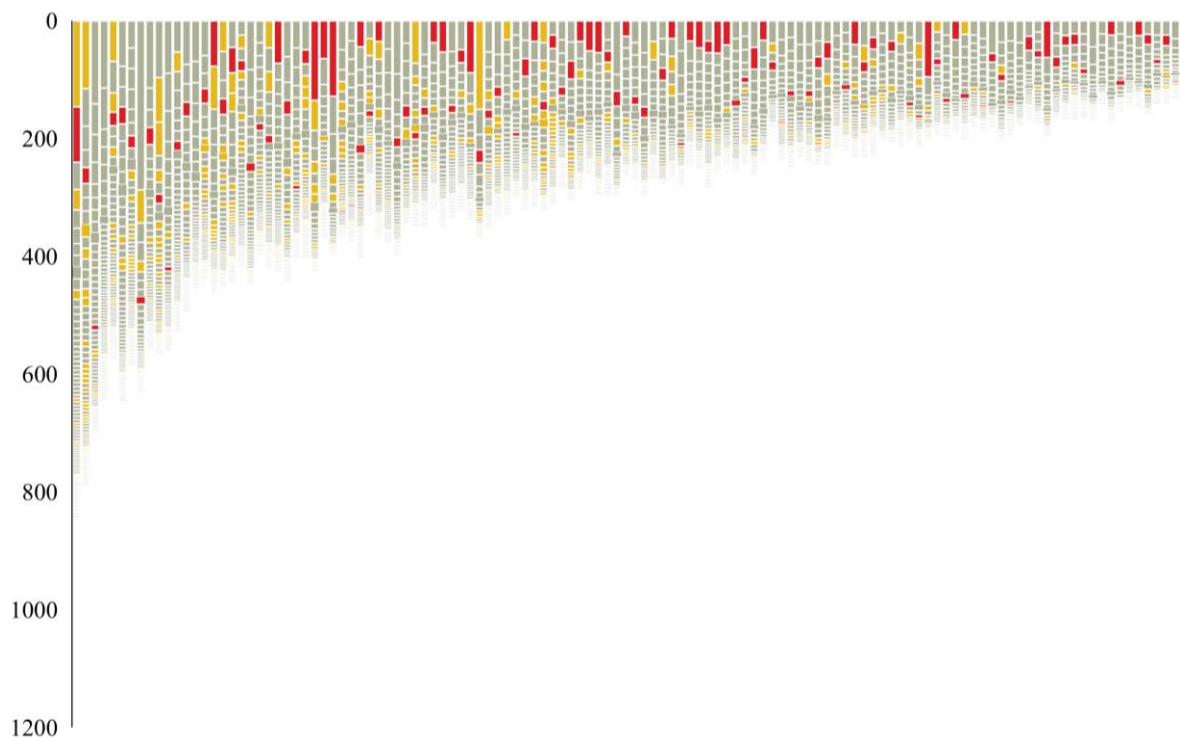


Figure 5 Rankings of publishing cities in all visiting cities to new papers

For the old papers, the rankings of the most of red data markers are not at the top of the bars, which means the publishing city is not the top visiting cities to the old papers in most situations, as Figure 6 shows. Compared with Figure 5, the area of the red and orange data markers shrinks a lot, especially the red data markers.

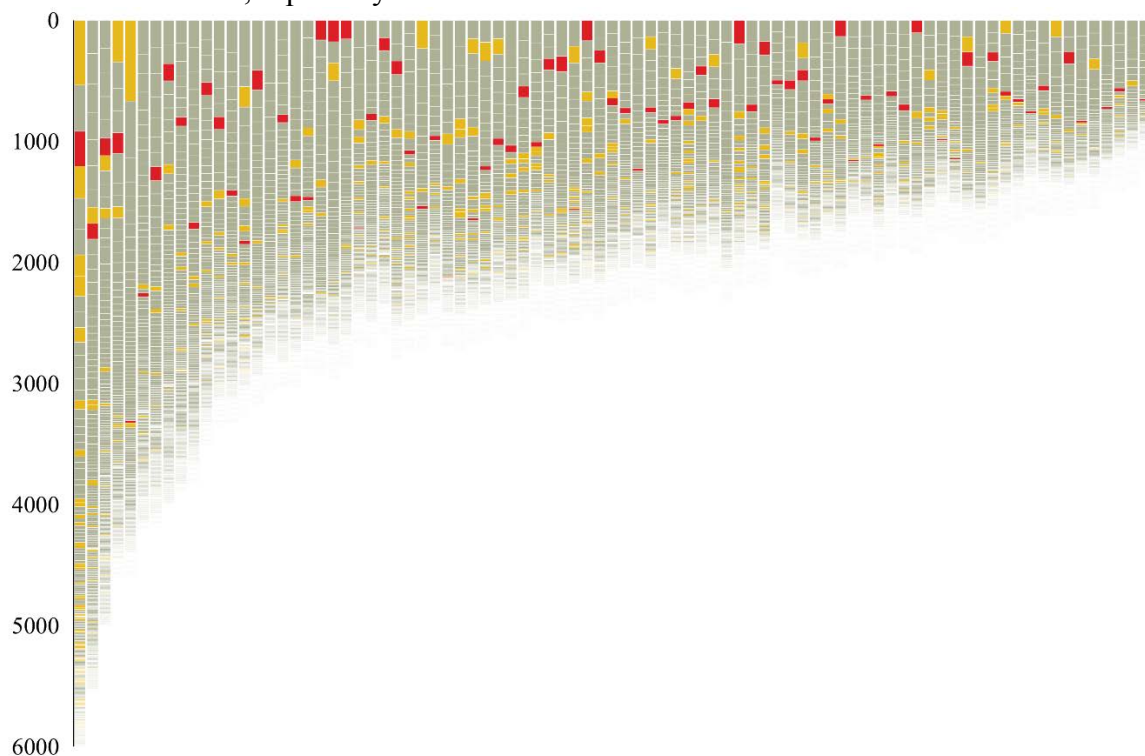


Figure 6 Rankings of publishing cities in all visiting cities to old papers

CONCLUSION AND DISCUSSION

In this study, we measure the innovation activities applying the usage data of scientific papers. Compared with the publishing data used by previous studies, article usage could reflect the innovation activities from another perspective with more detailed and real-time information. The geographical distribution at city level shows that most article visits are concentrated around major metropolitan areas and some high tech clusters. For the new papers and old papers, the visiting patterns from the geographical perspective are different. The visiting regions to new papers are dominated by the publishing cities, as time goes by, the visiting regions are diffused from the publishing cities to a broader area.

Besides the Open Access (OA) advantage to the authors, the Open Access movement also brings benefits to the developing countries. For many academics and universities in developing countries, journal subscriptions are often unaffordable. However, Open Access provides them the equality of access to the frontiers of scientific knowledge. In our study, we find that there are numerous visits from the developing countries, including many countries from Africa.

There are some limitations for this study. Because the data is harvested on one day and for one time, only the static results are given in this paper. If we want to know the real knowledge diffusion embodied in the scientific papers, it is necessary to track the dynamic diffusing process of the observations, which may need a long time period. We have been collecting the article usage data for the newly published papers since January, 2016. For each paper, the visiting data are updated daily, which makes it possible for us to know the whole diffusing process since the first day a paper published. We hope to dig deeper with this research topic and find out the patterns and mechanism of knowledge diffusing.

References

- Bornmann, L. & Leydesdorff, L. (2011). Which cities produce more excellent papers than can be expected? A new mapping approach, using Google Maps, based on statistical significance testing. *Journal of the American Society for Information Science and Technology*, 62(10), 1954-1962.
- Bornmann, L. & Waltman, L. (2011). The detection of “hot regions” in the geography of science—A visualization approach by using density maps. *Journal of Informetrics*, 5(4), 547-553.
- Bornmann, L., Leydesdorff, L., Walch-Solimena, C. & Ettl, C. (2011). Mapping excellence in the geography of science: An approach based on Scopus data. *Journal of Informetrics*, 5(4), 537-546.
- Brody, T., Harnad, S. & Carr, L. (2006). Earlier web usage statistics as predictors of later citation impact. *Journal of the American Society for Information Science & Technology*, 57(8), 1060-1072.
- Cottrill, C. A., Rogers, E. M. & Mills, T. (1989). Co-citation analysis of the scientific literature of innovation research traditions: diffusion of innovations and technology transfer. *Science Communication*, 11(2), 181-208.
- Davis, P. M. & Price, J. S. (2006). eJournal interface can influence usage statistics: implications for libraries, publishers, and Project COUNTER. *Journal of the American Society for Information Science and Technology*, 57(9), 1243-1248.
- Davis, P. M. & Solla, L. R. (2003). An IP - level analysis of usage statistics for electronic journals in chemistry: Making inferences about user behavior. *Journal of the American Society for Information Science and Technology*, 54(11), 1062-1068.

- Davis, P. M., Lewenstein, B. V., Simon, D. H., Booth, J. G. & Connolly, M. J. L. (2008). Open access publishing, article downloads, and citations: randomised controlled trial. *Bmj*, 337(7665), 1720-1720.
- Frenken, K., Hardeman, S. & Hoekman, J. (2009). Spatial scientometrics: towards a cumulative research program. *Journal of Informetrics*, 3(3), 222-232.
- Kaplan, N. R. & Nelson, M. L. (2000). Determining the publication impact of a digital library. *Journal of the American Society for Information Science*, 51(4), 324-339.
- Kotzeva, M., Brandmüller, T. & Önnersfors, Å. (2015). Eurostat regional yearbook 2015. Retrieved February 28, 2016 from http://ec.europa.eu/eurostat/statistics-explained/index.php/Research_and_innovation_statistics_at_regional_level.
- Kurtz, M. J. & Bollen, J. (2010). Usage bibliometrics. *Annual Review of Information Science & Technology*, 44(1), 1-64.
- Kurtz, M. J., Eichhorn, G., Accomazzi, A., Grant, C., Demleitner, M. & Murray, S. S., et al. (2005). The bibliometric properties of article readership information. *Journal of the American Society for Information Science & Technology*, 56(2), 111-128.
- Leydesdorff, L. & Persson, O. (2010). Mapping the geography of science: Distribution patterns and networks of relations among cities and institutes. *Journal of the American Society for Information Science and Technology*, 61(8), 1622-1634.
- Li, X., Chen, H., Huang, Z. & Roco, M. C. (2007). Patent citation network in nanotechnology (1976-2004). *Journal of Nanoparticle Research*, 9(3), 337-352.
- Liu, X., Jiang, S., Chen, H., Larson, C. A. & Roco, M. C. (2015). Modeling knowledge diffusion in scientific innovation networks: an institutional comparison between china and us with illustration for nanotechnology. *Scientometrics*, 105(3), 1-32.
- Luo, S., Du, Y., Liu, P., Xuan, Z. & Wang, Y. (2014). A study on coevolutionary dynamics of knowledge diffusion and social network structure. *Expert Systems with Applications*, 42(7), 3619-3633.
- Matthiessen, C. W., Schwarz, A. W. & Find, S. (2010). World cities of scientific knowledge: systems, networks and potential dynamics. An analysis based on bibliometric indicators. *Urban Studies*, 47(9), 1879-1897.
- Mazloumian, A., Helbing, D., Lozano, S., Light, R. P. & Börner, K. (2013). Global multi-level analysis of the 'Scientific Food Web'. *Scientific reports*, 3, 1167.
- Naumis, G. G. & Phillips, J. C. (2012). Diffusion of knowledge and globalization in the web of twentieth century science. *Physica A Statistical Mechanics & Its Applications*, 391(15), 3995-4003.
- Neff, M. W. & Corley, E. A. (2009). 35 years and 160,000 articles: a bibliometric exploration of the evolution of ecology. *Scientometrics*, 80(3), 657-682.
- Richard, V. N. (2010). Cities: building the best cities for science. *Nature*, 467(7318), 906-908.
- Shuai, X., Pepe, A. & Bollen, J. (2012). How the scientific community reacts to newly submitted preprints: article downloads, twitter mentions, and citations. *Plos One*, 7(11), e47523.
- Wang, X., Mao, W., Xu, S. & Zhang, C. (2014a). Usage history of scientific literature: nature metrics and metrics of nature publications. *Scientometrics*, 98(3), 1923-1933.
- Wang, X., Peng, L., Zhang, C., Xu, S., Wang, Z. & Wang, C. (2013b). Exploring scientists' working timetable: a global survey. *Journal of Informetrics*, 7(3), 665-675.
- Wang, X., Wang, Z. & Xu, S. (2013a). Tracing scientist's research trends realtimely. *Scientometrics*, 95(2), 717-729.

- Wang, X., Wang, Z., Mao, W. & Liu, C. (2014b). How far does scientific community look back?. *Journal of Informetrics*, 8(3), 562-568.
- Wang, X., Xu, S., Peng, L., Wang, Z., Wang, C., Zhang, C. & Wang, X. (2012). Exploring scientists' working timetable: Do scientists often work overtime?. *Journal of Informetrics*, 6(4), 655-660.
- Zhang, Q., Perra, N., Gonçalves, B., Ciulla, F. & Vespignani, A. (2013). Characterizing scientific production and consumption in Physics. *Scientific reports*, 3, 1640.



Public scientists contributing to local literary fiction. An exploratory analysis¹

Joaquín M. Azagra-Caro*, Anabel Fernández-Mesa and Nicolas Robinson-Garcia*

*jazagra@ingenio.upv.es; elrobin@ingenio.upv.es

INGENIO (CSIC-UPV), Universitat Politècnica de València, Valencia (Spain).

ABSTRACT

This paper presents an analysis of the role of scientists writing literary fiction in local publishing companies in the city of Valencia. As members of 'civic' universities, scientists are not only expected to produce new knowledge for the advancement of society, but also to account for their work and facilitate the public understanding of science. In this study we propose analysing their role as creators of literary fiction and the level of alignment of the fiction they produce with their academic activity. Here we present a small study focused on a single small publisher. In this first approach we find that local authors tend to be scientists, although there is no alignment between their literary work and their research.

INTRODUCTION

Public scientists (scientists only from now onwards), understood as a member of the teaching and/or research staff of a public university or a public research organization (including humanities and social sciences), benefit the academic community, industry and other social collectives through teaching and research. Most visible contributions are training skilled workforce, increasing the pool of knowledge and providing services to third parties. There are subtler contributions, like participation in unusual fora like diplomatic circles (Fährlich, 2015) or informal exchange of ideas with social agents that convert scientists into part of the cultural core of cities (Florida, 2005). Active involvement of scientists in culture is part of the richness of developed societies. Some voices in current debates on the evaluation of societal impact and the role of universities towards social development are claiming a refocus from a socioeconomic perspective to also including sociocultural benefits from academia (Goddard, 2009).

In this paper we will focus in one facet of cultural engagement; writing literary fiction. We depart from considering that a full understanding of scientists' contributions should not only account for their research work but also other academic activities. Eminent examples of scientists that have written literary fiction are Nobel Prize Winners like José Echegaray, Toni Morrison, John Maxwell Coetzee or popular authors like J.R.R. Tolkien and Philip Roth. The objective of this research is to analyse the contribution of scientists to literary fiction. We define as such genre/commercial fiction, skipping the debate of whether genre/commercial literature is fiction with literary value. Literary fiction refers to the narrative forms of literature, i.e. novels, short stories, plays; thus leaving aside poetry or literary essay. Such contribution can be from local and non-local authors. We will narrow our general objective to local activities, due to the interest in the engagement of scientist on this geographic

¹ Nicolas Robinson-Garcia is currently supported by a Juan de la Cierva-Formación Fellowship from the Spanish Ministry of Economy and Competitiveness.

dimension. Do local publishers include the literary work of scientists? Are works written by scientists more likely to be local than works not written by scientists?

Of course, scientists' literary fiction may not relate to their research. When there is a relation, literary fiction becomes a divulgation channel and, as such, a policy tool to increase knowledge diffusion and promote public understanding of science. This inquiry will also address a third research question: Do works written by local scientists present a higher degree of alignment between the contents of research and literary fiction?

DATA

Here we present preliminary result of an exploratory pilot study focused on a local publisher from the city of Valencia, Ediciones Contrabando and books published in 2015. By the date of the conference, we plan to expand to all publishers located in Valencia publishing literary works as defined above. Ediciones Contrabando was an interesting pilot study because it is a young company (born in 2012) and we have personal contacts with the chief editor, which is useful to explore in-depth issues that may arise when performing a larger scale analysis as well as work with a sample from which we can gain full understanding. Data was extracted from the Spanish national database of books published in Spain². This portal includes a database of Spanish publishers and a database of books. Ediciones Contrabando published in 2015 a total of 17 books. By manually checking online information about the books, we configured a valid sample of 9 fiction works (7 were non-fiction, 1 was not identifiable and 1 was a second edition of a book published the same year, which we excluded).

A well reported issue when working with monographs is the lack of address information of the authors (Gorraiz, Purnell & Glänzel, 2013), this makes it problematic when trying to identify the institutions behind such works or the background of the author. Here, the identification of scientists was done manually by checking in the Internet and searching for information on the author's affiliation.

METHOD

In order to measure the contribution of scientists to literary fiction, our proposal is to identify which proportion P of edited books of literary fiction is authored by scientists. We will answer our first research question by estimating the model:

$$P_{ijkl} = f(L_j, Z_{ijkl})$$

Here, P is the probability that the edition i of book j by author k , published by company k , has an academic author. L represents whether the author is from the same city of the publishing company. Z is a vector of control characteristics.

For academic authors, we will measure the degree of alignment D between their research and their fiction work. We will consider three possible degrees: 0 (none), 1 (formal alignment, i.e. the action occurs in a scientific setting), 2 (content alignment, i.e. the plot has to do with the scientist's research lines). For further analysis, we will estimate the model:

$$D_{ijkl} = f(L_j, Z_{ijkl})$$

² <http://www.mecd.gob.es/cultura-mecd/areas-cultura/libro/bases-de-datos-del-isbn.html>

PRELIMINARY RESULTS

Ediciones Contrabando published 9 fiction books in 2015. 3 had scientific authors, i.e. 33%. This is a first quantification of the contribution of public scientists to literary fiction. At first sight, probably an overestimation that further research and enlargement of the dataset will nuance, but also a manifestation the phenomenon is present.

Local authors were 7, i.e. 78%. This is coherent with the intuition that young, small editorial companies rely on geographic proximity to nurture their portfolios. The 3 scientific authors were local, which implies that the localisation between editor and authors increases the probability of finding scientific authors. By the day of the conference, we will test this via econometric models with a larger number of observations.

The degree of alignment between the research lines of the scientific authors and their fiction is null in the three cases. Hence, in this pilot study, we would observe that scientific authors do not diffuse academic knowledge through their literary works, and that co-location with the editor does not play a role.

DISCUSSION AND FURTHER RESEARCH

As members of publicly-funded institutions, scientists have a central role on the creation of new knowledge in order to facilitate social progress. Lately, a key concern for policy makers is to learn about the societal contribution of scientists for social development. In this sense, contributions are normally understood in terms of research activity and output, and not in terms of the implications of having an academic population embedded in a local environment. Here we build on the creative class of Florida (2005) where it is suggested that the inclusion of a certain group of individuals enhances local development and enrichment. We propose to analyse their role in literary fiction, specifically we are interested in analysing the level of alignment between their work as academics and their oeuvre, and see if they use it as a tool for facilitating the public understanding of science. As ‘civic’ universities where researchers are viewed as workers for the public good, activities such as literary publishing can serve as powerful tools for popularizing science (Goddard, 2009).

Here we propose a first approach based on local publishers in a given city. This way we can control variables such as the propensity of scientific authors to publish with local companies, the relation between having research institutions and a larger number of literary scientists, etc. However, we still need to overcome important limitations. Here we show a very small sample of books from a small publishing company. The difficulty to automatically identify the affiliation of authors remains a considerable shortcoming for working with large datasets which could provide us a better understanding and insight of the role of scientists in local literary publishing. Finally, we must point out other issues that will be solved in the future such as the identification of genre and other classical problems when working with monographs (Torres-Salinas et al., 2014) such as translations, new editions, edited books of short stories, etc.

REFERENCES

Fährnich, B. (2015). Science diplomacy: Investigating the perspective of scholars on politics–science collaboration in international affairs. *Public Understanding of Science*, 0963662515616552. <http://doi.org/10.1177/0963662515616552>

Florida, R. (2005). *Cities and the creative class*. Routledge.

Goddard, J. (2009). *Re-inventing the civic university*. NESTA.

Gorraiz, J., Purnell, P. J., & Glänzel, W. (2013). Opportunities for and limitations of the Book Citation Index. *Journal of the American Society for Information Science and Technology*, 64(7), 1388–1398

Torres-Salinas, D., Robinson-Garcia, N., Cabezas-Clavijo, Á., & Jiménez-Contreras, E. (2014). Analyzing the citation characteristics of books: Edited books, book series and publisher types in the Book Citation Index. *Scientometrics*, 98(3), 2113-2127.



Does collaboration facilitate the performance of enterprise innovation?

Qi Lv^{*}, Donghua Zhu^{*}, Ying Huang^{*}, Liliana Mitkova^{**}, Xuefeng Wang^{*}, Gizem Ogsuz^{**}

^{*}performerlv@163.com; zhudh111@163.com; huangying_work@126.com; wxf5122@bit.edu.cn; School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China

^{**}Liliana.Mitkova@u-pem.fr; gizemoguz@yahoo.com; Univ Paris Est Marne Vallee, F-77420 Champs Sur Marne, France

ABSTRACT

Open innovation has been an important way for companies acquiring knowledge resources to achieve rapid development externally, especially those aiming to obtain key technology capabilities. In this paper, we propose a method to analyse the innovation performance of the open innovation practices with two aspects of indicators: degree of collaboration and innovation performance. Using a sample of sixteen leading firms in graphene all over the world, empirical results identify that open innovation conducts to the intensification of patents production, patent citations and researches' diversity. The results show that width, depth and diversity of collaboration all has a positive effect on innovation performance.

KEYWORDS

Open innovation, Collaboration, Innovation performance, Graphene, Patents analysis.

INTRODUCTION

Since 1970, China transformed its economy from planned closed to open and integrated to market economies. During this transformation, China targeted to become an innovation country itself. Recent researches show that the Chinese firms are using different instruments such as technology licences (Liu et al, 2006), collaboration with universities and R&D institutes (Chen and Qu, 2003) and alliances (Duysters et al, 2009) for increasing their innovation performance. The aim of any firms is to optimize the performance and profit using external resources *via* knowledge exchanges and cooperation with various partners. In this context, firms set up partnerships to widen their internal knowledge, gain access to complementary technological resources, insulate from environmental uncertainty, access new market and preserve technological leaderships (Prahalad and Hamel, 1990, Burgers et al., 1993, Mortehan, 2004). In the open model these practices are based on different forms of inter-organizational relationships such as R&D collaboration, or licensing-in and licensing-out developing their absorptive and disruptive capabilities (Bianchi et al., 2011; Jeong et al., 2012;).

Among different open innovation instruments, we are asking the importance of the collaboration on company's innovation performance. In order to respond to this question, this research contributes to the open innovation literature by proposing an integrated method to evaluate the performance of the collaborative open processes at the example of the graphene industry.

THEORETICAL BACKGROUND

Open innovation

Open innovation is generally studied from two forms: inbound and outbound flows of the knowledge. Inbound open innovation is the firms' acquisition of the external knowledge to use within the company and to increase their innovative capacity (Chesbrough, 2006). Nowadays reaching external knowledge has been accepted as an important factor of successful innovations (Rothwell, 1994). Besides, outbound open innovation is the activity of sharing the internal knowledge with the outsiders. According to open innovation scholars, knowledge landscape is widely distributed by the efforts of universities, technological institutions, national labs and small-medium sized enterprises (Enkel, E., Oliver, G. & Chesbrough H., 2009).

Innovation Performance

Innovation performance, in its broader definition, is a process, as it encompasses outcomes from the conception of an idea all the way to the introduction of a product or process into a market (Meglio, 2009). It reflects long-term gains through the invention of new process- and product-related technologies. These new technologies can eventually lead to improved profitability for companies if they are transformed into actual innovations (Finkelstein and Cooper, 2009).

Firms can measure the efficiency of the collaborations' innovation performance by the output, which in that case the intellectual property or the new product creation. In the open innovation theory, the evaluation of the collaboration performance is generally made by taking into account the R&D expenditure, the number of new products and number of patents (Ahuja and Katila, 2001).

In conclusion, the literature proposes the variety of studies addressing collaborative innovation performance in terms of operational indicators, variables, temporal orientation, methods, and data. In order to contribute to the building of a comprehensive and complex evaluation of innovation performance of collaborative innovation, the objective of this paper is to suggest a method guided by the following driving questions: how to involve more quantitative methods to measure innovation performance of technological collaboration?

METHODOLOGY

Indicators

Degree of collaboration

We define collaboration as a case in which two companies are both assignees in a patent. So the

measurement of Degree of collaboration from patents data lies on co-assignee analysis. As is known to all, an assignee in a patent is similar to the institution of authors in an article. Undoubtedly, co-occurrence matrix is a detailed and powerful expression of a firm's collaboration situation in a certain period of time. From that matrix, we calculate three different indicators to measure the factor.

(1) Width of Collaboration (WC) is defined as how many partner enterprises have been involved in collaborations.

$$WC = TN \quad (1)$$

In Equation (1), TN is the total number of entities with which one enterprise has cooperated.

(2) Depth of Collaboration (DC) is defined as the times of co-operations of one firm in a certain year. The more a company applies for patents with partners together, the higher it's DC will be.

$$DC = AC \quad (2)$$

In Equation (2), AC is the aggregated count of patents one enterprise applies for in a year in which assignees also include other entities.

(3) Diversity Index (DI) is a coefficient that measures the degree of dispersion (Huang et al., 2015). If one company impartially cooperates with all its partners we can conclude that it has a relative high diversity index.

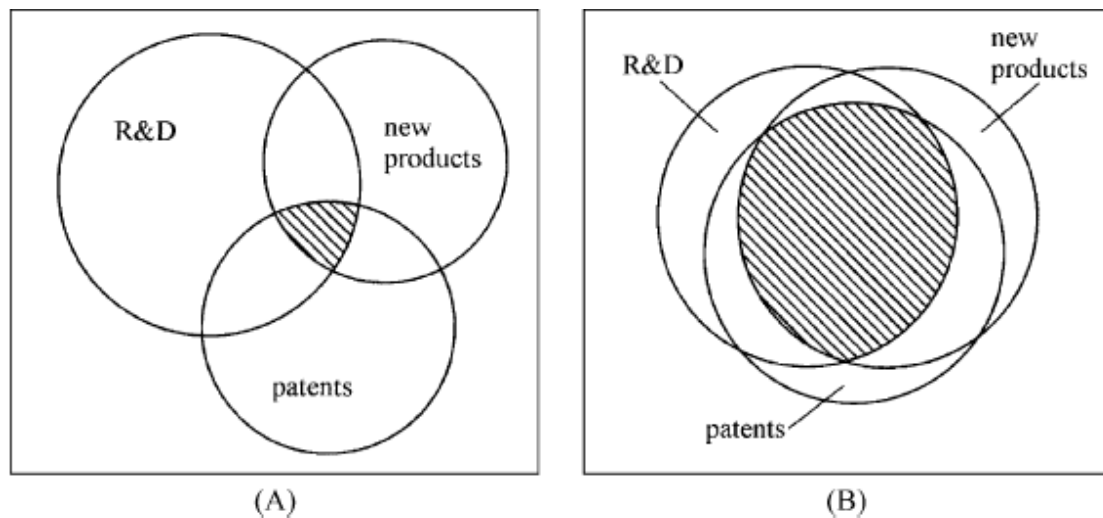
$$DI = 1 - \frac{\sum_{i=1}^k Q_i^2}{Q^2} \quad (3)$$

In Equation (3), $i=1,2,\dots,k$. It represents the serial number of a firm's partners (including itself), Q_i equals to count of patents it applies for through collaboration with partner i . Q represents the total multi-organizational patents so $Q = \sum_{i=1}^k Q_i$.

Innovation performance

As mentioned before, available measures contain R&D inputs, patents and new product announcements. The overlap between the three aspects is highly relative with industry. As Hagedoorn et al.(2003) put it, when the industry has high R&D intensity, high patenting intensity and high ratio for new product introduction, the intersection will be large (as shown in Figure 1 (B)).

Figure 1. Intersection of three indicators in different industries



Obviously, graphene is a rapid-changing industry with high R&D intensity, we can pick out one aspect and believe that it is representative of all the three. We define three indicators as follows:

- (1) Patents Count (PC) is the total amount of patents of one firm published in a year. It reflects the quantity of patent research activity.
- (2) Patents Citation Count (PCC) is the total amount of citations (being cited) of a set of patents published in the same year. It reflects the quality of research activity. Considering that patents published early may naturally obtain more citations, we use $PCC/(2016-t)$ as a relative fair measure, where t is publication year and ranges from 2001 to 2015.
- (3) Specialization Score (SS) measures the degree of specialization (Porter et al., 2008),

$$SS = \frac{\sum(f_i * f_j * \cos(IPC_i, IPC_j))}{\sum(f_i * f_j)}$$

In the equation, f_i is the proportion of patents assigned to each IPC class i (you can use the amount or frequency, either is okay). Cosine is used as a measure of similarity between the two IPC classes. The vectors are all calculated from the citation data in European Patent Office (EPO).

Model Specification

We assume that an enterprise's collaboration has positive effect on its innovation performance. So we try to use regressive model to describe that effect. Since the time span is 2001 to 2015, we use panel data regressive model to analyze the trend.

The regression model is fixed effects model; this kind of model can remove individual effects and fully use panel data. Using dependent variable PC as an example, we define models like:

$$PC_{i,t} = \beta_0 + \beta_1 WC_{i,t} + \beta_2 DC_{i,t} + \beta_3 DI_{i,t} + \beta_4 Scale_{i,t} + a_i + u_{i,t}$$

In this equation, we use three indicators to measure collaboration as independent variables: WC, DC and DI. i is the serial number of organizations, t is the serial number of year. Scale, which is measured by a company's employee number, performs as a control variable. a_i represents the individual features of each firm, for example, enterprise culture, spirit of innovation and organizational structure. The corner mark doesn't contain time variable t , meaning that these features are stationary.

RESULTS

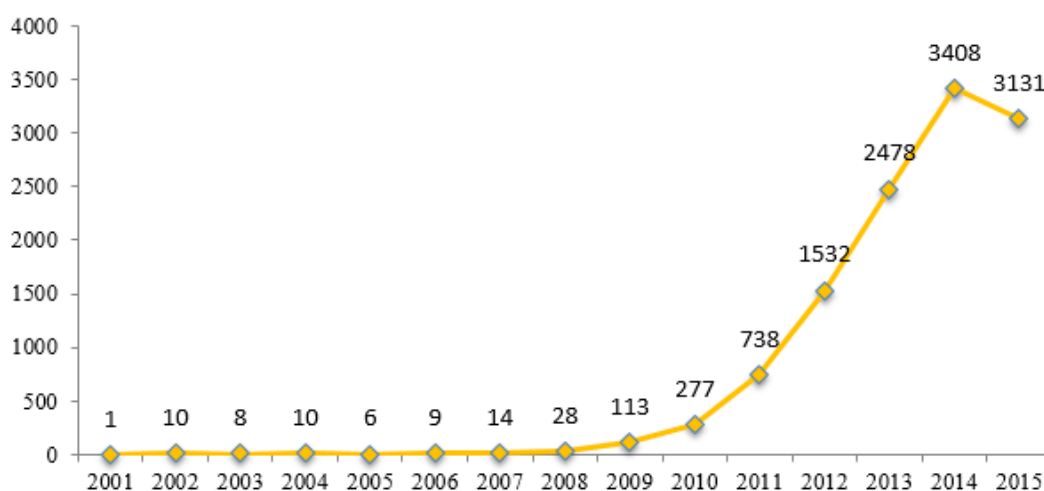
We choose Thomson Innovation (TI) (<https://www.thomsoninnovation.com>) to retrieve patent data on graphene all over the world. Data was downloaded on January 5, 2016. The search strategy is “TI= (Graphene) AND DP>= (20010101) AND DP<= (20151231)”. 11763 records of patents were obtained.

Descriptive statistics

Time series of patent count

As is shown in Figure 2, since 2004 (production of the first single-layered graphene), publication of patents on graphene started to grow exponentially and reached its peak in 2014.

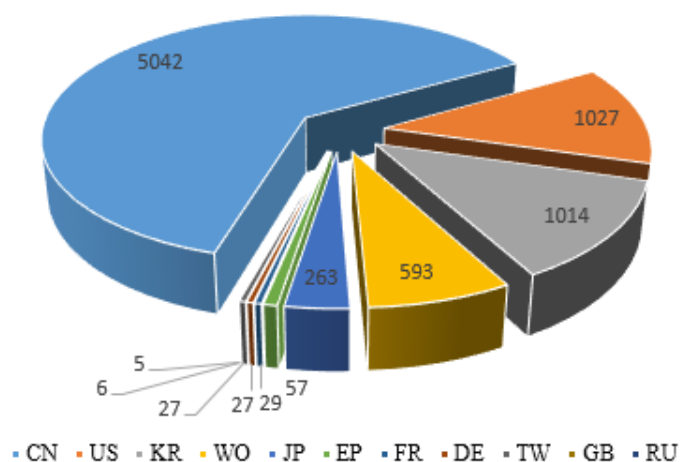
Figure 2. Patent count on Graphene around the world from 2001 to 2015



Assignees' distribution

Raw data contains large amount of reputation in assignees' name. With the help of Vantage Point, we finally found 4013 assignees during 2001-2015.

Figure 3. Assignees' distribution among countries



We select the top 500 assignees in patent count to analyze the distribution among countries. From Figure 3, we see that China, United State and Korea are the most animated market countries in graphene researches.

Since we are interested in the dynamic of enterprises, we list top 16 firms below.

Table 1. Top 16 Enterprises in patent publication

Rank	Patent Assignees	Country	Patents
1	SHENZHEN OCEANS KING LIGHTING SCI & TECH (OKLS)	China	353
2	SAMSUNG ELECTRONICS CO LTD (SMSU)	Korea	335
3	INT BUSINESS MACHINES CORP (IBMC)	United State	136
4	LG ELECTRONICS INC (GLDS)	Korea	85
5	CHONGQING MOXI TECHNOLOGY CO LTD (CHON-N)	China	71
6	HON HAI PRECISION IND CO LTD (HONH)	China	58
7	TOSHIBA KK (TOKE)	Japan	58
8	UNIST ACAD-IND RES CORP (UNIS-N)	Korea	44
9	DOKURITSU GYOSEI HOJIN SANGYO GIJUTSU SO (NIIT)	Japan	42
10	NANOTEK INSTR INC (NANO-N)	United State	39
11	EMPIRE TECHNOLOGY DEV LLC (EMPI-N)	United State	38
12	FUJITSU LTD (FUIT)	Japan	36
13	POSCO (POSC)	Korea	33
14	GRAPHENE SQUARE INC (GRAP-N)	Korea	31
15	SEKISUI CHEM IND CO LTD (SEKI)	Japan	29
16	VORBECK MATERIALS CORP (VORB-N)	United State	29

From Table 1 we can conclude that Asian enterprises have a dominant position in graphene study. Korean companies are the world leader in this field, Japan and China are competitive too.

Regression Analysis

The panel data we choose contains the top 16 firms' collaboration and patents information in year 2008-2015. Firstly, some variables are different in range: DC, WC, PC, PCC and Scale are all natural numbers, DI and SS range from 0 to 1. We take the logarithm of DC, WC and Scale to reduce the range. For example, we use $NDC=\ln(DC+1)$ to replace DC, $NWC=\ln(WC+1)$, $NScale=\ln(Scale+1)$.

Adjustments of Models

On the one hand, we found that there is a strong correlation between WC and DC. On a high R^2 (0.7921) standard we got that $DC=1.56*WC+0.96$. In order to avoid multicollinearity, we should use one of them instead of two as independent variables.

Table 2 . Regression output of WC to DC

Corr(u_i,Xb)=0		Prob>chi2=0.000		
DC	Coef	Std. Err	z	p> z
WC	1.56	0.074	21.02	0.000
cons	0.96	0.830	1.16	0.248
R²(over-all)= 0.7921				

On the other hand, we should consider time lag variables (Ahuja and Katila, 2001). Theoretically, co-operation activities usually take a firm some time (may be several years) to absorb the knowledge or skills learnt during certain collaboration. We can set additional variables in equation as $DC_{i,t-k}$, $DI_{i,t-k}$ to represent the effects of efforts in history. Here k is the value of time lag. Consider the short time-span of data, we use k=1 in order to avoid producing much missing values.

To sum up, the equation was adjusted like:

$$PC_{i,t} = \beta_0 + \beta_1 NDC_{i,t} + \beta_2 NDC_{i,t-1} + \beta_3 DI_{i,t} + \beta_4 DI_{i,t-1} + \beta_5 NScale + a_i + u_{i,t}$$

Statistical software is STATA12.0 was used to estimate the parameters.

Estimation Results

1. Relation between collaboration and patent count

Table 3. Regression output of DC, DI, Scale to PC

Wald chi2(5) = 157.10		Prob > chi2 = 0.0000		
PC	Coef	Std.Err	z	p> z
NDC	2.89	0.37	8.30	0.000
lagNDC	0.87	0.09	-1.32	0.185
DI	0.10	0.06	-4.05	0.000
lagDI	2.25	0.98	1.85	0.064
NScale	1.40	0.16	2.99	0.003
cons	0.20	0.09	-3.63	0.000

It can be seen in Table 3 that all indicators of collaboration have positive effect on PC, DC and one year lag variable of DI show bigger effects. It means that finding more partners for research, the more patents they tend to acquire. Meanwhile, Scale of company also has a significant positive effect on PC. We can explain that a large company usually has more researchers and assets for doing research, leading to a higher possibility of breakthrough in technology.

2. Relation between collaboration and patent citation count

Table 4. Regression output of DC, DI, Scale to PCC

Corr(u_i,Xb)=-0.0162		Prob>F=0.0017		
	Coef	Std.Err	t	p> t
NPCC				
NDC	4.42	1.96	2.26	0.026
lagNDC	-8.95	2.11	-4.23	0.000
DI	-11.20	7.36	-1.52	0.131
lagDI	25.67	7.98	3.22	0.002
NScale	1.50	1.63	0.92	0.360
cons	2.48	5.73	0.43	0.666
R²(With-in)=0.5794				

1 year lag variable of DI still has positive effect on PCC. It is DC not 1 year lag variable of it that has a positive effect on PCC, showing that depth of collaboration may contribute to patents quality more quickly.

3. Relation between collaboration and specialization score

Table 5. Regression output of DC, DI, Scale to SS

Corr(u_i,Xb)=-0.3019		Prob>F=0.0000		
	Coef	Std.Err	t	p> t
Inss				
NDC	-0.08	0.04	-2.28	0.025
lagNDC	0.01	0.04	0.25	0.8
DI	-0.17	0.14	-1.23	0.222
lagDI	-0.20	0.15	-1.33	0.186
NScale	-0.09	0.03	-2.98	0.004
cons	1.13	0.11	10.42	0.000
R²(With-in)= 0.5315				

Unfortunately, three independent variables are not significant even at the significance level of 0.1, which may illustrate that specification score is affected by some extra variables that we don't consider. It worth noting that DC and Scale both show a negative effect on SS: with 1% growth of co-operation diversity, the specialization degree may decrease 0.08%. This can be explained that co-operating with partners in various fields can expand a firm's horizon, so it may develop more complicated technology that involves more IPC classes, leading to a low specialization degree.

CONCLUSIONS AND DISCUSSIONS

The study identified the impact of collaboration on innovation performance from a bibliometric point of view. The results enrich existing research dedicated on the open innovation putting the accent on the criteria as one of the most important aspect on the decision-making. We

systematically introduced the indicators in two aspects: WC, DC and DI to measure degree of collaboration, PC, PCC and SS to measure innovation performance.

The industry of the case study is graphene. We choose 16 representative enterprises from 2008-2015 to analyze their patents data. Multiple regression of fixed-effects is used to deal with panel data. The results show that collaboration's width and depth are highly correlative. Depth and diversity of collaboration both have positive effects on patent count and patent citations. DI has a negative impact on specialization degree. It is noticeable that the changes of DI have approximately one year's lag to affect three dependent variables. Finally, high standard of collaborations tend to broaden a company's horizon and lead to a more diverse scope of IPC in patents published.

This study contributes in different ways to the open innovation literature. From a theoretical point of view, the proposed model adds to previous researches dedicated to the innovation performance by putting the accent on the technological aspect of innovation performance.

This study has some limitations. First, the definition of collaboration is narrow. The collaboration between enterprises should contain R&D contract, M&A, licenses and other factors. To measure these factors, traditional methods contain investigations and interviews. Second, the two variables WC and DC are tested to be correlative, so we have to abandon one independent variable. So we are still considering introducing several new independent variables. Third, graphene is an emerging industry so the time span is too short. Some companies' indicators from 2008-2010 are almost all zero, causing bad effects on the regression results. Even the panel data is still not sufficient to support a convincing enough regression results. Similar study should better be carried out in a more mature field with relative long time span.

REFERENCES

- Ahuja, G., & Katila, R. (2001). Technological acquisitions and the innovation performance of acquiring firms: A longitudinal study. *Strategic management journal*, 22(3), 197-220.
- Bianchi, M., Cavaliere, A., Chiaroni, D., Frattini, F., & Chiesa, V. (2011). Organisational modes for Open Innovation in the bio-pharmaceutical industry: An exploratory analysis. *Technovation*, 31(1), 22-33.
- Burgers, W. P., Hill, C. W., & Kim, W. C. (1993). A theory of global strategic alliances: The case of the global auto industry. *Strategic management journal*, 14(6), 419-432.
- Chen, J., & Qu, W. G. (2003). A new technological learning in China. *Technovation*, 23(11), 861-867.
- Chesbrough, H. W. (2006). *Open innovation: The new imperative for creating and profiting from technology*. Harvard Business Press.
- Duysters, G., Jacob, J., Lemmens, C., & Jintian, Y. (2009). Internationalization and technological catching up of emerging multinationals: a comparative case study of

- China's Haier group. *Industrial and corporate change*, 18(2), 325-349.
- Enkel, E., Gassmann, O., & Chesbrough, H. (2009). Open R&D and open innovation: exploring the phenomenon. *R&d Management*, 39(4), 311-316.
- Finkelstein, S., & Cooper, C. L. (2010). *Advances in mergers and acquisitions*. Emerald Group Publishing.
- Hagedoorn, J., & Cloudt, M. (2003). Measuring innovative performance: is there an advantage in using multiple indicators?. *Research policy*, 32(8), 1365-1379.
- Hamel, G., & Prahalad, C. K. (1990). The core competence of the corporation. *Harvard business review*, 68(3), 79-91.
- Huang, Y., Ma, J., Porter, A. L., Kwon, S., & Zhu, D. (2015). Analyzing collaboration networks and developmental patterns of nano-enabled drug delivery (NEDD) for brain cancer. *Beilstein journal of nanotechnology*, 6(1), 1666-1676.
- Jeong, S., Lee, S., & Kim, Y. (2013). Licensing versus selling in transactions for exploiting patented technological knowledge assets in the markets for technology. *The Journal of Technology Transfer*, 38(3), 251-272.
- Liu, J. J., Qian, J. Y., & Chen, J. (2006). Technological learning and firm-level technological capability building: analytical framework and evidence from Chinese manufacturing firms. *International Journal of Technology Management*, 36(1-3), 190-208.
- Meglio, O. (2009). Measuring performance in technology-driven M&As: Insights from a literature. *Advances in Mergers and Acquisitions*, 8, 103.
- Meglio, O., & Risberg, A. (2010). Mergers and acquisitions—Time for a methodological rejuvenation of the field?. *Scandinavian Journal of Management*, 26(1), 87-95.
- Mortehan, O. (2004). The role of firms collaborative agreements in the information technology industry transformation. *Technology Analysis & Strategic Management*, 16(1), 53-71(19).
- Porter, A. L., Roessner, D. J., & Heberger, A. E. (2008). How interdisciplinary is a given body of research?. *Research Evaluation*, 17(4), 273-282.
- Rothwell, R. (1994). Towards the fifth-generation innovation process. *International marketing review*, 11(1), 7-31.



Structural Analysis of Redundancy Influence of Local Regions in Renewable Energy R&D Projects in Europe

Jaso Larruscain-Sarasola^{*}, Rosa Río-Belver^{*} and Gaizka Garechana^{**}

^{*} jaso.larruscain@ehu.eus; rosamaria.rio@ehu.eus

Department of Industrial and Project Engineering, University of The Basque Country (Upv/Ehu), Nieves Cano 12, Vitoria-Gasteiz, 01006 (Spain)

^{**} gaizka.garechana@ehu.eus

Research Department of Industrial Engineering, University of the Basque Country (UPV/EHU), C/Elcano 21, Bilbao, 48008 (Spain)

ABSTRACT

The increase of the percentage of Renewable Energy (RE) in the European energy mix to 20% by 2020 caused the boost of a new multi-discipline industrial sector, becoming Knowledge and Innovation networks at local level (nuts3) an industrial hallmark. Smart quantitative indicators related to knowledge and geographical territory help to understand the coherence and usefulness of the public investment in the sector. For this purpose, combining Text Data Mining (TDM), Boolean Search Logic, cleaning process and geolocalization techniques to Community Research and Development Information Service database (CORDIS), 1,118 projects and 10,104 partners were identified for 2000-2013. The usefulness of Social Network Smart indicators based on traditional approach on centrality and new structural holes-based focusing on the position of influence of each actor in the network were tested. The results show that not always the relationships between projects are done efficiently (avoiding redundant links) and individual partners are unaware of their influence in the overall network when choosing a project partner. Policy makers should take into account the global picture of the network when configuring R&D project grants. The use of Structural Holes measures is useful to generate partners/local regions rankings according to their influence and the efficiency of the relationships.

INTRODUCTION

The objective to increase the percentage of Renewable Energy (RE) in the European energy mix to 20% by 2020 has caused the boost of a new multi-discipline industrial sector including not so well-developed technologies such as wave, tidal and small wind energies, and the creation of new local organizational structures (Marques and Fuinhas, 2012). Knowledge and Innovation networks at local level in this emergent RE sector is an industrial hallmark.

Smart quantitative indicators related to knowledge and geographical territory help to understand the coherence and usefulness of the public investment in the sector.

The interrelations between partners (firms, universities and research centers) allow the transfer of knowledge which leads to a process of development of the innovation (Löfsten and Lindelöf, 2005) linked to the proximity concept. Since each actor does not have the vision of the overall network structure, policy makers should identify which organizations or local

regions (at Nuts3 level) participating in R&D projects have better network position to avoid redundant investment or isolated collaborations.

This research aims to describe the usefulness of Social Network Smart indicators based on Structural Holes theory which analyses how to create non-redundant R&D project collaboration networks in local level (geographical Nuts3) in RE sector, identifying the key players and their relationships.

METHOD

Since European Union finances most R&D collaborative technological projects in RE in Europe, Community Research and Development Information Service database (CORDIS) was chosen as a source of information which includes 163,664 actors and 97,992 projects financed by European Union Budget since 1981.

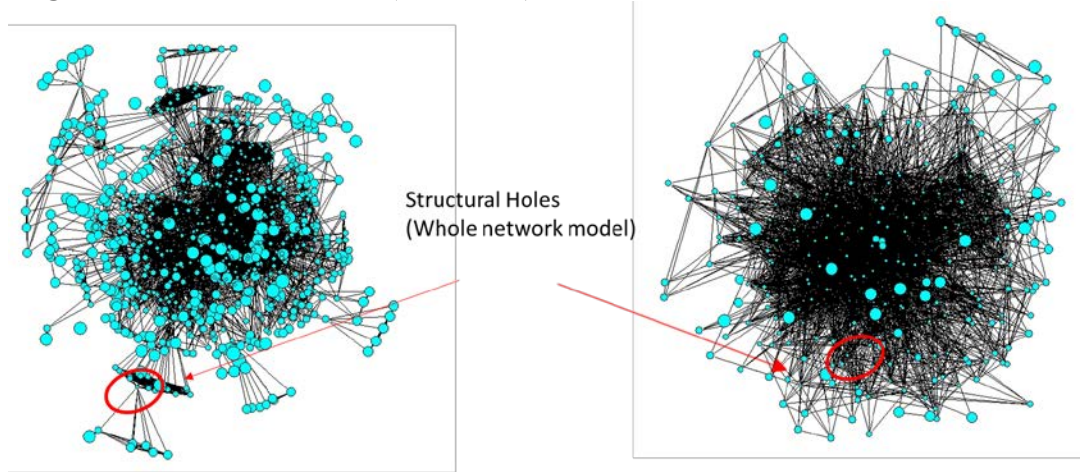
Firstly, combining Text Data Mining (TDM), “Boolean Search Logic” with “Specific Area Classification Sources” for RE technological area (Kostoff 2006) applied to “Keywords”, “Title” and “Activity” fields, a new data base was created with 6,703 R&D RE projects for Wind, Solar, Biomass, Geothermal and Tidal/Wave energy areas. 1,118 projects and 10,104 partners were filtered for 2000-2013. A cleaning process to obtain an accurate data with Vantage Point software and geolocalization techniques with GPSVisualizer package were applied.

Secondly, Social Network Analysis (SNA) is used as an empirical method to obtain the dynamic and structural properties of the relationships between partners (Batallas and Yassine 2006). Two main approaches were studied: traditional approach based on centrality (betweenness, degree and closeness) and new structural holes-based focusing on the position of influence of each actor in the network.

RESULTS AND DISCUSSION

The new approach is suitable for gauging the degree of redundancy of an organization's contacts in collaboration network and to measure efficiency and effectiveness of its connections (Kang and Park, 2013; Borgatti et al., 2002). This enables the study on how to economize on the number of ties required to access unique information and earn control acting as brokers between disconnected organizations (Baum et al., 2003). Constraints indicator was considered in ego and whole-network model, using Pajek and Ucinet software (Figure 1).

Figure 1: Wind R&D network (2000-2013). Partners-Partners and Nuts3-Nuts3 network



Both approaches show similar general position for partners on the top of the ranking (Table 1). However, for the rest of the partners in the ranking both approaches differ considerably. To understand the difference in detail, Structural Hole based approach was studied and some key results were emerged.

On the one hand, since Constraint indicator measures a high level of knowledge diversity, high level of radical innovation and influence in ties (Kang and Park, 2013) separating groups of organizations with low grade of common information, partners situated on the top of the ranking seem to have similar influence in each of the RE sectors, but mainly in Wind, Solar and Biomass sectors. Moreover, partners with high influence in Sea and Geothermic sectors do not have or have low influence in the rest of the RE sectors.

On the other hand, local regions in Europe (Nuts3) have similar influence in each of the sectors (Table 2). Additionally, our study concluded that specially in the Sea sector, local region influence is correlated to the number of projects carried out in there.

Table 1: Centrality vs. Structural Holes-based approaches for partners in RE (2000-2013).

Number	Partner	Centrality	Struct. Hole
1	Fraunhofer E.V.	0.2975	0.0088
2	CIEMAT	0.2264	0.0112
3	Centre National de la Recherche Scientifique	0.2174	0.0118
4	Centre for Renewable Energy Sources	0.2157	0.0124
5	University of Stuttgart	0.1968	0.0126
6	EDF S.A.	0.1825	0.0145
7	Consiglio Nazionale delle Ricerche	0.1745	0.0154
8	Centre for Research and technology Hellas	0.1745	0.0157
9	National Technical University of Athens	0.1851	0.0158
10	Commisariat Energies Alternatives	0.1824	0.0176
11	Technical University of Delft	0.1767	0.0183
12	Teknologian Tutkimuskeskus VTT	0.1619	0.0183
13	Danmarks Tekniske University	0.2010	0.0188
14	Nederlandse Organisatie Voor Toegepast Natuurwetenschappelijk Onderzoek TNO	0.1600	0.0194
15	Stichting Energieonderzoek Centrum Nederland	0.1817	0.0195

Table 2: Structural Holes-based indicators ranking of local regions (2000-2013).

NUTS3	RE	Wind	Sea	Solar	Biomass	Geotherm
DE212	1.0000	0.9510	0.9620	1.0000	0.9692	0.1622
ITE43	0.9984	0.9878	0.5823	0.9957	0.9846	0.9730
FI181	0.9969	0.9673	0.0000	0.9399	1.0000	0.0000
FR102	0.9953	0.8980	0.9241	0.9914	0.9949	0.9459
GR300	0.9938	0.9796	0.9873	0.9785	0.9923	0.9595
NL333	0.9922	0.9265	0.7975	0.9742	0.9744	0.9324
ES300	0.9906	0.9755	0.2025	0.9979	0.9795	0.8649
DE111	0.9891	0.9143	0.0000	0.9936	0.9436	0.0000
AT130	0.9875	0.6449	0.0000	0.9807	0.9897	0.0000
ES511	0.9860	1.0000	0.0000	0.9850	0.8538	0.0000
ITC45	0.9844	0.9918	0.0000	0.9893	0.9564	0.6892
AT221	0.9828	0.3918	0.0000	0.8991	0.9821	0.3649
BE100	0.9813	0.9959	0.7089	0.9871	0.9974	0.4595
SE121	0.9797	0.7878	0.8861	0.8906	0.9667	0.0000
HU101	0.9782	0.3592	0.0380	0.9678	0.8872	0.9865

CONCLUSIONS

Applying Social Network Smart indicators based on Structural Holes theory at local level (geographical Nuts3), some key conclusions were emerged. Firstly, not always the relationships between projects are done efficiently (avoiding redundant links). Secondly, individual partners are unaware of their influence in the overall network when choosing a project partner. Policy makers should take into account the global picture of the network when configuring R&D project grants. The use of Structural Holes measures is useful to generate partners/local regions rankings according to their influence in other partners and the efficiency of the relationships.

REFERENCES

- Batallas, D.A. & Yassine, A.A. (2006). Information leaders in product development organizational networks: social network analysis of the design structure matrix. *IEEE Transactions on Engineering Management*, 53(4), 570–582.
- Baum, J.A.C., Shipilov, A.V. & Rowley, T.J. (2003). Where do small worlds come from? *Industrial and Corporate Change*, 12(4), 697–725.
- Borgatti, S.P., Everett, M.G. & Freeman, L.C. (2002). *Ucinet for Windows: Software for Social Network Analysis MA: Analytic Technologies*. London: Harvard.
- Kang, M.J., & Park, J. (2013). Analysis of the partnership network in the clean development mechanism. *Energy Policy*, 52(0), 543-553.
- Kostoff, R.N. (2006). Systematic acceleration of radical discovery and innovation in science and technology. *Technol. Forecast. Soc. Change*, 73(8), 923–936.

Löfsten, H. & Lindelöf, P., (2005). R&D networks and product innovation patterns-academic and non-academic new technology-based firms on Science Parks. *Technovation*, 25(9), 1025–1037.

Marques, A.C. & Fuinhas, J.A. (2012). Is renewable energy effective in promoting growth? *Energy Policy*, 46, 434-442.



The discrepancy of patent citation behavior between examiners and inventors: a citation network analysis¹

Ying Huang^{*}, Donghua Zhu^{*}, Qi Lv^{*}, Alan L. Porter^{**}, and Xuefeng Wang^{*}

^{*}huangying_work@126.com; zhudh111@163.com; performerlv@163.com; wxf5122@bit.edu.cn
School of Management and Economics, Beijing Institute of Technology, Beijing 100081 (China)

^{**}alan.porter@isye.gatech.edu
School of Public Policy, Georgia Institute of Technology, Atlanta, GA 30332 (USA)
Search Technology, Inc., Atlanta, GA 30092 (USA)

INTRODUCTION

Patent citation analysis is increasingly applied in the past ten years, because it has intrinsic advantages in treating impacts of scientific or technical advances quantitatively. Developing from purely citing-and-cited analyses to patent co-citation, bibliographic coupling and self-citation, patent citation analysis has become a powerful method used for a variety of different objectives. However, a crucial factor we need to care about is that patent citations can be included by the applicant and also can be added by the patent examiner responsible for judging the degree of novelty of the patent. In other words, there are two kinds of reference citations in patents: applicant citation and examiner citation.

In this paper, we attempt to explore the discrepancy of patent citation behavior between examiners and inventors in the scope of patent citation network analyses in a particular technology area. In detail, we choose photovoltaic cells (PV) as our target technologies, and retrieve related USPTO granted patent data. Based on coverage statistics of applicant and examiner citations, we analyze some network parameter indicators of three kinds of patent citation networks. Then, we introduce specialization scores to probe the similarity of technology field distribution of patents included in the above three modes.

SEARCH TERMS AND DATA SOURCE

We collected USPTO patent grant data from Thomson Innovation (<https://www.thomsoninnovation.com>). We chose PV as our example and the search term we used is: CPC=((Y02E10/54*)). In this paper, publication date is applied, considering the patent time lag, and the retrieval period we choose is between January 1, 2001 and December 31, 2015. The reason we choose 2001 as the start is that USPTO provides information on the source of patent citations since 2001. Ultimately, 5899 records are obtained, and the nine technologies of PV and the numbers of patents under study are shown in Table 1.

¹ This work was supported by US National Science Foundation (Award #1527370 - "Forecasting Innovation Pathways of Big Data & Analytics").

Table 1. Search Results on Nine Sub-technologies of Photovoltaic Cells (PV)

CPC Code	Sub-technologies	Short Name	Records
Y02E 10/541	CuInSe2 material PV cells	CuInSe2	661
Y02E 10/542	Dye sensitized solar cells	DSSCs	728
Y02E 10/543	Solar cells from Group II-VI materials	Group II-VI	230
Y02E 10/544	Solar cells from Group III-V materials	Group III-V	753
Y02E 10/545	Microcrystalline silicon PV cells	Microcrystalline	172
Y02E 10/546	Polycrystalline silicon PV cells	Polycrystalline	243
Y02E 10/547	Monocrystalline silicon PV cells	Monocrystalline	1399
Y02E 10/548	Amorphous silicon PV cells	Amorphous	501
Y02E 10/549	Organic PV cells	Organic PV	2190

Note: Some patents are tagged in more than a single category.

RESULTS

Citation Coverage Statistics

The impact of examiners in patent citations is significant because they can add, retain or even delete original inventors' citations, and the amount of US examiner citations is usually larger than the number of inventor citations (Meyer 2000). We explored these patterns in more detail by examining the distribution of these shares over sub-technologies, shown as Table2. There are three aspects of conclusions we can obtain: First, most PV patents have citations added by applicant or examiner, and more than half the patents have both types of citation. In the sub-fields of CuInSe2 and Microcrystalline, the no-citation coverage is zero. Even in DSSCs, 96.8% of the patents have at least one citation. Second, both kinds of citation distribute in approximately 90% of patents. However, examiner citation is more frequent than applicant citation, in 7 of these 9 sub-technologies, except for DSSCs and Organic PV. Third, the amount of applicant citation is visibly larger than that of examiner citation. On average, the citations added by applicants is 5 times larger than those added by examiners. Especially in Groups II-VI, the former is 7 times larger than the later.

Table 2. Propensity of Inventor and Examiner Citation for Nine Sub-technologies

Sub-technologies	Non-Citation	Applicant Citation		Examiner Citation	
	Coverage	Coverage	Average Citation	Coverage	Average Citation
CuInSe2	0.000	0.837	35.546	0.887	6.318
DSSCs	0.032	0.828	14.357	0.810	4.400
Group II-VI	0.009	0.865	36.226	0.891	4.552
Group III-V	0.003	0.822	28.983	0.854	4.697
Microcrystalline	0.000	0.884	17.942	0.901	4.895
Polycrystalline	0.004	0.868	26.449	0.877	5.132
Monocrystalline	0.001	0.873	28.391	0.881	4.931
Amorphous	0.002	0.896	17.856	0.898	5.082
Organic PV	0.015	0.866	23.759	0.770	4.095

Network Parameter Analysis

The pattern of technology transfer and knowledge flow can be generally divided into two types: interior citation behavior and exterior citation. In this paper, we mainly focus on the interior citation that indicates the citation relationship presented among the patents in the targeted technology, rather than beyond the technology field. We apply centrality analysis to better compare the different structures of three patent citation networks: citation by applicant, citation by examiner, and citation by both applicant and examiner. Comparative results are shown as Table 3.

In terms of four centrality indicators, the average values of patents owning applicant citation are higher than the ones having examiner citation in degree-centrality, closeness-centrality, and betweenness-centrality, especially in degree centrality. Such results may partly be due to the number of citations offered by applicants being much larger than by examiners in the PV field. In addition, the differences become more obvious if we take the networks built by applicant vs. examiner citations into consideration. In term of Standard Deviation, the betweenness-centrality of (applicant & examiner) citation networks is much greater than those limited to either applicant's and examiner's, but the closeness-centrality and eigenvector centrality among them are almost similar.

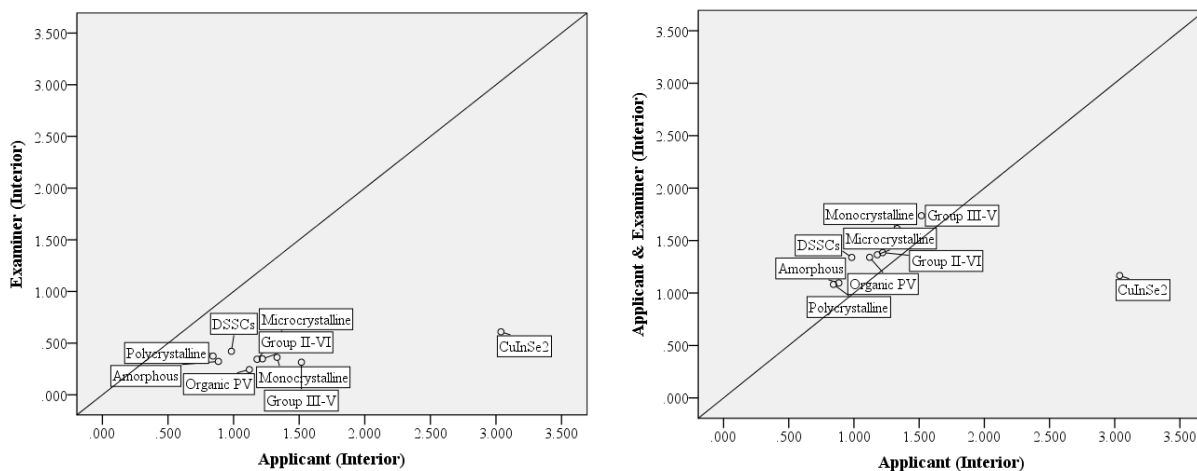
Table 3. Centrality Analysis of Three Kinds of Patent Citation Network

	Applicant & Examiner (Interior)		Applicant (Interior)		Examiner (Interior)	
	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
Degree centrality	4.065	8.745	3.125	7.515	0.717	1.637
Closeness centrality	0.255	0.406	0.198	0.378	0.155	0.348
Betweenness centrality	3.892	26.58	1.594	14.007	0.245	1.919
Eigenvector centrality	0.008	0.031	0.006	0.028	0.007	0.033

Furthermore, the network of patents connected by citations provides a representation of the innovation process (Erdi et al. 2013). To further analyze the network parameter, we use three indicators reflecting main properties of patent citation networks to conduct pairwise comparison for nine sub-technologies in PV. These are shown as Figure 1 (average weighted degree), Figure 2 (average path length), and Figure 3 (average clustering coefficient).

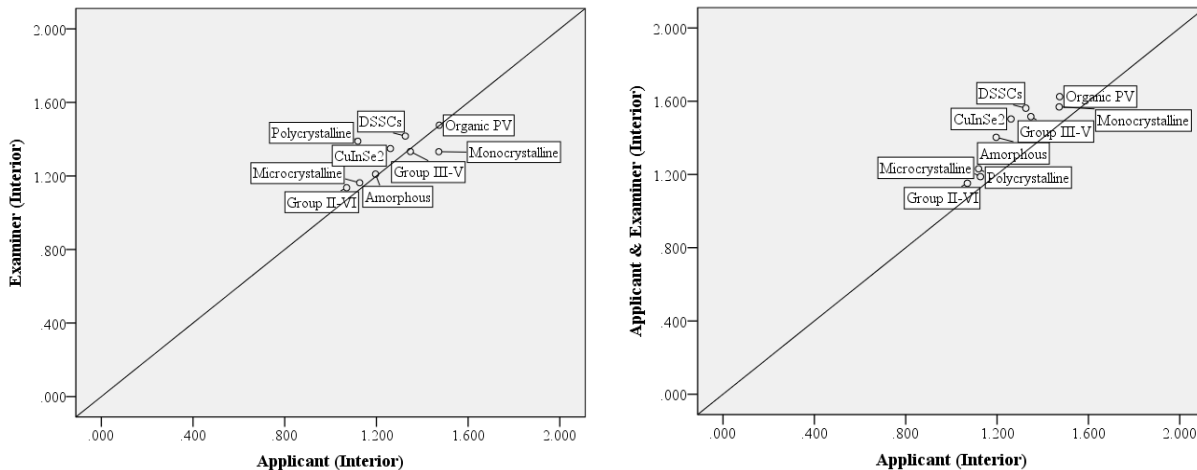
The average weighted degree expresses the average degree of a node in a weighted network, and it is equal to half of the average degree in a binary network. Its value also relies on the number of edges connected to it, and closely relates to the density of a network, including in-degree (citing other nodes) and out-degree (cited by other patents). Compared with examiner citation, applicant citation generates a much stronger citation relationship among interior patents. When we merge examiner citations with applicant citations, the merged citation network presents more citation relationship than individual applicant citation network, except for the CuInSe₂, which shows higher average weighted degree in applicant's only citations.

Figure 1. Comparisons of Average Weighted Degree of Applicant vs. Examiner (left) and Applicant vs. (Applicant & Examiner) (right)



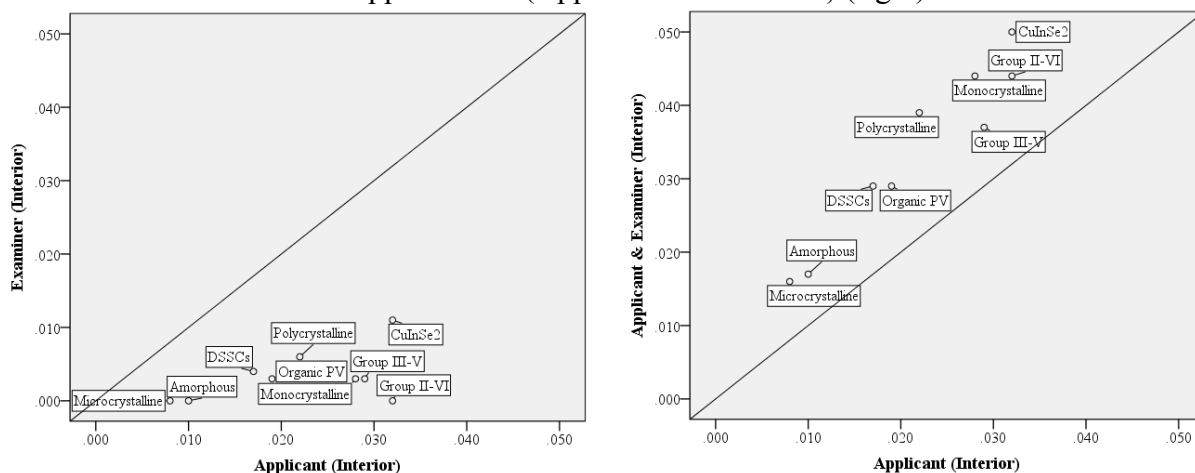
The average path length is calculated by the shortest path between all pairs of nodes, and shows the number of steps it takes to get from one patent of the citation network to another. Figure 2 reveals that examiners' citation behavior has less influence in connecting these nodes that are not originally linked when taking just the applicant citation into consideration.

Figure 2. Comparisons of Average Path Length of Applicant vs. Examiner (left) and Applicant vs. (Applicant & Examiner) (right)



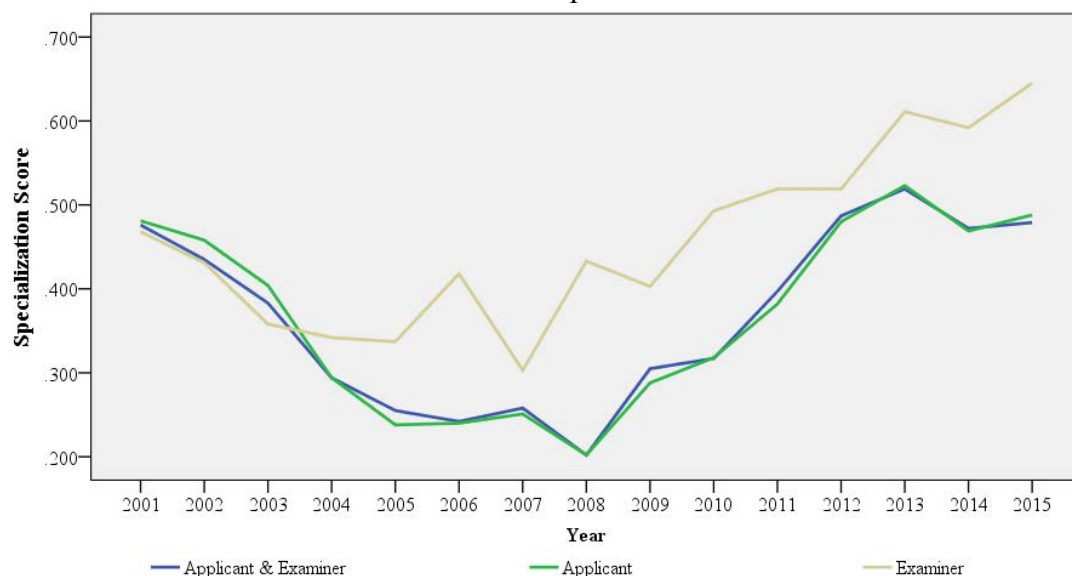
The clustering coefficient indicates how nodes are embedded in their neighborhood and shows the degree of aggregation in a network. Along with the mean shortest path, it can present any “small-world” effect. The average gives an overall indication of clustering in the network. Contrasted with the examiner citation networks, the application citation networks have a much higher value of average clustering coefficient; that is, the patents having applicant citation behavior are more likely to reflect co-citation activities. However, the value of the average clustering coefficient rises once we consider both applicant citation behavior and examiner citation behavior. In other words, the merged citation network more easily forms a close component, and different nodes have smaller average distance.

Figure 3. Comparisons of Average Clustering Coefficient of Applicant vs. Examiner (left) and Applicant vs. (Applicant & Examiner) (right)



Patent Specialization Analysis

The specialization score is originally applied to evaluate the diversity of Web of Science Subject Categories (SCs) assigned to the journal in which the publication (or set of publications) is printed (Porter et al. 2008). In this paper, we used International Patent Classifications (IPCs) in place of SCs to gauge the degree of concentration of technology fields of a series of patents. Figure 4 traces the difference between the specialization degree of examiner citation and applicant citation by year. In the beginning years (2001 to 2003), the patents among applicant citation networks and those in examiner citation networks have similar technology spans. After 2004, on one hand, the patents included in examiner citation networks are likely to concentrate on a relatively more narrow technological field than in applicant citation networks. Thus, a higher specialization score would result; on the other hand, the patents in applicant citation networks and (applicant & examiner) citation networks reveal similarly concentrated technology fields. The number of patents contained in these two networks shows subtle differences.

Figure 4. Technology Specialization Degree of Patents Included in Three Types of Interior Giant Components

CONCLUSIONS AND DISCUSSION

In this paper, we analyze a particular technology area (photovoltaic cells), rather than the full patent dataset in USPTO or EPO. We apply citation coverage statistical analysis, network analysis, and patent specialization analysis to explore the discrepancies of patent citation behavior between examiners and inventors in the scope of resulting patent citation networks. We have obtained four conclusions. First, most patents have at least one or more citations, but the share of USPTO examiner citation is lower than that of applicant citation. This point goes against many results of previous studies. Second, adding examiner citation to applicant citation enhances performance in constructing a more strongly connected network, when we come to the average clustering coefficient. Third, the patents included in an examiner citation network are more specialized in relatively narrow technological fields. This supports the view that examiners may increase the quality of prior art searching since they always have real experience and knowledge of the pertinent art.

REFERENCES

- Érdi, P., Makovi, K., Somogyvári, Z., Strandburg, K., Tobochnik, J., Volf, P., & Zalányi, L. (2013). Prediction of emerging technologies based on analysis of the US patent citation network. *Scientometrics*, 95(1), 225-242.
- Meyer, M. (2000). What is special about patent citations? Differences between scientific and patent citations. *Scientometrics*, 49(1), 93-123.
- Porter, A. L., Roessner, D. J., & Heberger, A. E. (2008). How interdisciplinary is a given body of research? *Research Evaluation*, 17(4), 273-282.



How Does Technology Transfer from Universities to Market in China? An Empirical Analysis Based on Invention Patent Assignment ¹

Yang Yang*, Kun Ding**, Chunbo Zhang**, Xiaoling Sun** and Zhigang Hu**

*yangyang0477@mail.dlut.edu.cn

WISE Lab, Dalian University of Technology, Dalian, 116024 (China)

**dingk@dlut.edu.cn; zhangcb@mail.dlut.edu.cn; xlsun@dlut.edu.cn; huzhigang@dlu.edu.cn

WISE Lab, Dalian University of Technology, Dalian, 116024 (China)

ABSTRACT

Technology transfer of universities and research commercialization have been paid extensive close attention by Chinese government in recent years and a series of policies have been launched to facilitate these activities. In this study, a novel approach is proposed to explore the patent diffusion from universities to market in order to study the technology transfer modes. We combine text mining, network analysis and especially data visualization as an integrated tool to analyse patent assignment. The patent data used in this paper are all from SIPO (State Intellectual Property Office) Database and these patents belong to 985-project universities in China. The total number of the invention patents in our research is 105770 and the number of patents that include patent assignment is 3177. Our method firstly extracts both sides that participate in a patent transaction from the unstructured patent data. Then a transfer path for each patent is generated based on the date when patent assignee changes. With that, we build a patent-diffusion network from different paths and visualize it with some special methods. Finally, this study summarizes seven technology transfer modes among Chinese universities semi-automatically according to the nature of the entities, paths of high frequency and nodes of high score. Universities in China mainly concentrate on Mode 1, Mode 2 and Mode 3.

INTRODUCTION

As an important part of regional and national innovation system, university is not only a producer of knowledge and talents, but also the important source of science and technology. At present, more comprehensive attention has been taken to the importance of research commercialization and technology transfer of universities. In the early 1980, the Bayh-Dole Act was enacted in the U.S. to permit a university or a private sector to pursue ownership of an invention in preference to the government from federally funded research and this act provided incentive and effective policy for the research commercialization. Similar initiatives are underway in other countries (Abreu & Grinevich, 2013; Rasmussen, 2008). At the same time, lots of universities have started to facilitate technological entrepreneurship and try to transform themselves from research universities to entrepreneurial universities (Kalar & Antoncic, 2015).

A lot of researches focus on transfer channels between universities and industry, such as patenting, spin-off creation, licensing, consultancy, contract research, joint research and

¹ This work was supported by the project of “National Natural Science Foundation of China” (61272370, 71473028)

training (Bekker & Freitas, 2008; Landry et al., 2010; D'Este & Patel, 2007). Grimaldi et al. (2013) referred to these activities as “academic entrepreneurship” due to the fact that the objective of such efforts is commercialization of innovations developed by academic scientists. Some other studies are conducted to investigate the influence factors and determinants of university technology transfer including internal intellectual property regulations, cultivation of talents, individual characteristics, organizational factors, institutional environment, government policy and regional context (Baldini et al., 2006; Swamidass & Vulasa, 2009; Huang et al., 2011; Grimm & Jaenicke, 2012). Harmon et al. (1997) mapped the technology transfer process based on 23 technologies from university of Minnesota, and investigated the effects on the transfer process with the firm size, the stage of company growth and informal networking arrangements. Markman et al. (2005) built a framework to reveal the complex relationships among university technology transfer office (UTTO) structures, strategies, new venture formation and business incubation with grounded theory. Grimpe and Fier (2010) focused on the informal university technology transfer based on the interactions between university scientists and industry personnel. However, the studies based on the indicator of patent assignment are rather rare. Our study fills this gap by analysing patent transaction data of universities in China.

Patent assignment is a very special way of interaction between universities and enterprises. It can not only help the inventors and universities realize the commercial value of new knowledge, but also can solve the technical problems for enterprises to promote the innovation of products and upgrade industries (Rao et al., 2011). Patent assignment in universities is highly valued by academia, industry and government. Since 1985 the first invention patent authorized, invention patent authorization has amounted to 1.53 million by the date of 24/12/2014 in China. However, the rate of technology transfer in China is only about 10%, which is far less than other developed countries (Zhang, 2013). A similar situation which also appears in the universities of China, the utilization of academic patents is at a low level (Luan et al., 2010). It is necessary and urgent to explore how technology (patent) transfers from universities to market and summarize universal technology transfer modes as a guide. Song et al. (2011) tried to use questionnaire to find out the patent transfer modes in Chinese universities. In his research, 157 colleges and universities, 668 enterprises were investigated. In our study, we use patent transaction data of 985-project universities to summarize technology transfer modes in Chinese universities.

DATA AND METHODOLOGY

Each university has two types of patent data as shown in figure 1. Type 1 describes the basic information of current patent such as id, title, abstract, type, applicant, legal status, etc. Type 2 describes all transaction data including patent-id, transaction-type, transaction-date and the details of transaction. 25 kinds of transaction-type have been set in SIPO patent database, but only 2 kinds are related to research topic, one is patent assignment (Kind 17) and the other is effect, change and cancellation in contract for patent license (Kind 19). In this paper, we focus on patent assignment (Kind 17) for analysis due to that the patent license only involves two sides, licensor and licensee, and in our study, we also want to find the middle nodes in the diffusion of patent.

Figure 1. The structure of patent data

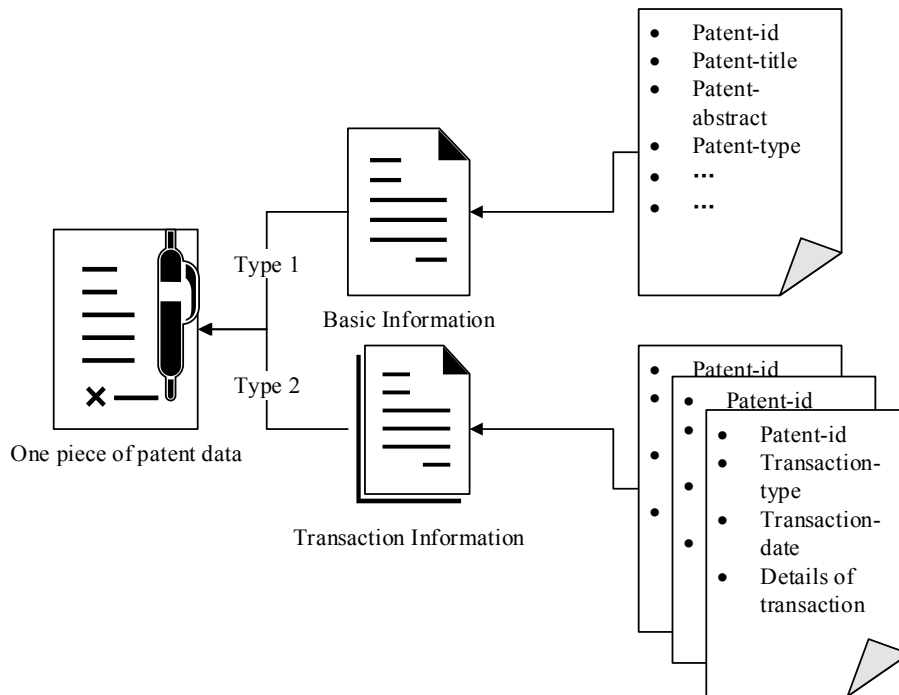
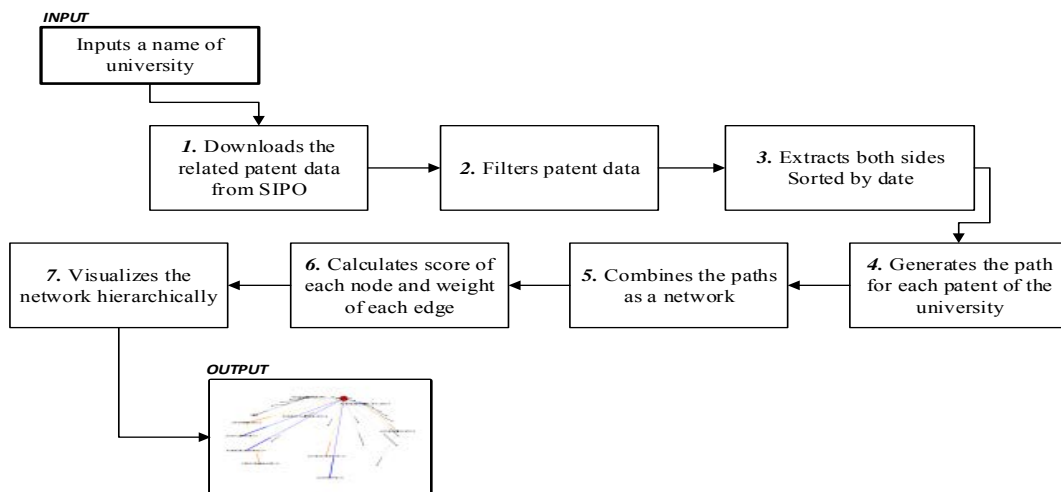


Figure 2. Framework of patent-transfer system



As shown in figure 2, there are nine parts with different functions in our patent-transfer system. Firstly, a name string of university is entered, if local JSON files (structured text) of this university exist, a graph of hierarchical network will be generated as the output. Otherwise, if local data of the university does not exist, the system will download and filter the patent data on SIPO website automatically, then patent data will be parsed and integrated into JSON files. In order to describe in a more logical way, seven parts are finally combined into four big parts as shown in Table 1.

Table 1 Four important parts in whole system

<i>New Part</i>	<i>The numbers of old part</i>
Data Crawling	1
Data Filtering	2
Data Processing	3,4,5,6
Data Visualization	7

Data Crawling

Basic information and transaction information about one patent can be found in SIPO website, <http://epub.sipo.gov.cn/gjcx.jsp> and <http://epub.sipo.gov.cn/flzt.jsp>. Web crawler technology is applied in patent-transfer system instead of manual downloading. Basic data and transaction data of each patent are combined into one JSON file based on the same patent-id.

Data Filtering

As mentioned above in data part, invention patents which are authorized are chosen as our research corpus, but co-applicant patent is filtered out and with the transaction status, kind 17 is selected. So, there are two tasks in this part. One task is to filter the patent data with two or more applicants, and the other one is to filter the patent data by selected rules in options.

Data Processing

- *Step 1: Extracts both sides (two entities)*

Regular expression is used in this part. In order to validate the method, two persons are chosen to extract the entities manually from 100 pieces of transaction data. Compared with manual results, the accuracy of our regular expression template is 100%. It is of great significant in this part, and with that perfect work, following missions can be done. Each patent data is parsed into a collection of triples. [(Entity A, Entity B, Date), (Entity C, Entity D, Date), (Entity B, Entity C, Date), etc.]. A triple like (a, b, 2014.1.1), it means the patent application right transfers from a to b in 1/1/2014

- *Step 2: Combines entities and generates the propagation path*

In this mission, the collection of triples for each patent data is sorted by the transaction date. Entities are combined by the rule of head and tail echo, most of the paths can be generated easily as $A \rightarrow B \rightarrow C \rightarrow D$. But sometimes, because of the mistakes of SIPO patent database, some special processing must be carried out to solve problems. For example, the name of entity B is ‘李四有限责任公司’, and there is also an entity C named ‘李四有效责任公司’, while the collection of triples is [(A, B, Date), (C, D, Date)], because of the input mistake of one character, B and C becomes the different entities, but in fact they are the same. So, it is impossible to form a path in this situation by the rule of head tail echo. In order to make computers recognize that B and C are the same, Levenshtein distance (Levenshtein, 1996) is used in our research to solve this problem. The levenshtein distance of two entities is less than or equal 1 if two entities are the same.

- *Step 3: Calculate score of each node and weight of each edge*

If we consider the different entities as different nodes, the different relationships as paths, it is not hard to form a directed graph which shows the diffusion of patent. As a directed graph, in-degree and out-degree are the basic attributions of each node and a node with high betweenness centrality has a large influence on the transfer of items through the network. In our study out-degree and betweenness centrality of each node are used to calculate the final score of the node (excluding the root node) and a normalized function is used to make the

values between 0 and 1. As shown in formula 1, x_a is the normalized out-degree value of node a . $out_degree(a)$ is the out-degree value of node a . Min and Max represent the minimum and maximum in the out-degree values of nodes. Formula 2 is same as formula 1, and it is used to calculate the normalized betweenness value of node a .

$$x_a = \frac{out_degree(a) - Min}{Max - Min} \quad (1)$$

$$y_a = \frac{betweenness(a) - Min}{Max - Min} \quad (2)$$

Finally, a linear equation is used to combine x_a and y_a as shown in formula 3 and in our study variable c has an empirical value 0.8. It is worth to point out that in our experiment the root node of a university or an institute of the university is very special because that its out-degree value and betweenness value are very high. That will impact the normalized function and make the score of other nodes very low. So, in our research the second largest value is selected to be Max instead of the value of the root node.

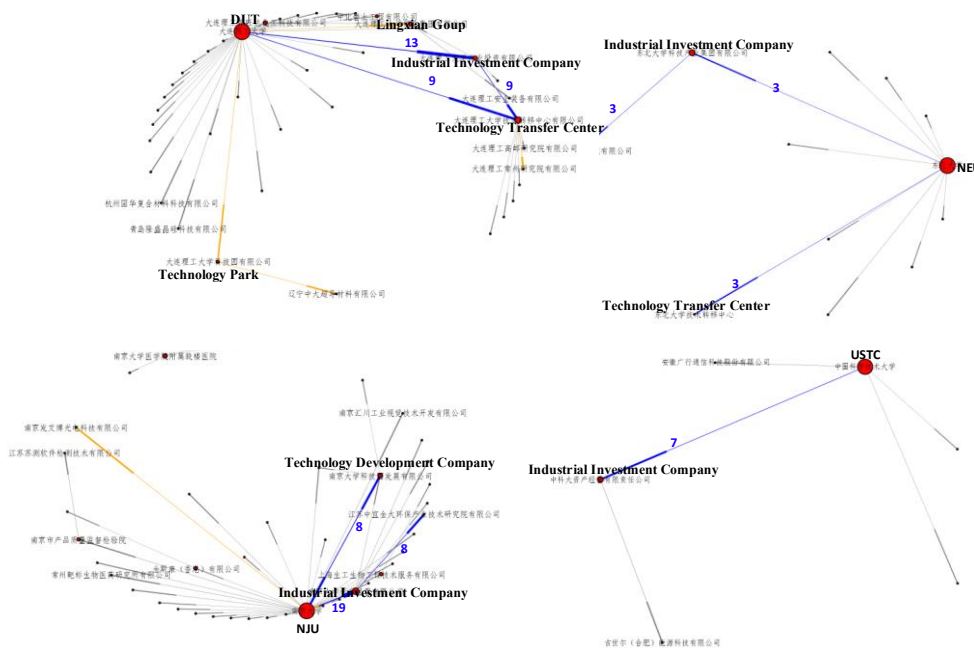
$$Score_a = cx_a + (1 - c)y_a \quad (3)$$

Data Visualization

In order to visualize the network hierarchically and mark the important nodes and paths obviously, $Score_a$ and the weight of edge is used. As shown in formula 4, R_a is the radius of node a (excluding the root node) in the directed graph. R_1 and R_2 are two variables that adjust according to the actual significant in the network. In our research, the value of R_1 is 50, the value of R_2 is 4. Because of the reason that mentioned above, the radius of root node is a constant which is 200 in this paper.

$$R_a = Score_a \times R_1 + R_2 \quad (4)$$

Figure 3. Examples for Visualization



As shown in figure 3, different colours of paths are shown in the figure. Blue is selected for the top 5% edges and Orange is selected for top 5%-10% edges. The edges are sorted by the weights and the thick head of each edge is the direction of technology transfer. Numbers beside the blue edges are their weights and represent the number of patent assignments. When the number of patent assignments is larger, the edges are thicker.

RESULTS

As shown in figure 4, the authorized patents of Zhejiang University and Tsinghua University go ahead in China, but the transfer rate of patent is low. Meanwhile, the patent transfer rate of NPU (Northwestern Polytechnical University) and SEU (University of Southeast) is high, they have a common feature and both have patent transfer relationship with many enterprises in Hai'an County, Nantong, Jiangsu Province. Actually enterprises in Hai'an County have a strong cooperation with many colleges and universities in China, this condition is inseparable with local management and policy.

Table 2. Seven kinds of transfer mode

<i>Number</i>	<i>Transfer Mode</i>
1	Direct Transfer of Patents
2	Asset Management Unit of University
3	University-Enterprise Cooperation in R&D
4	Local Government-University Cooperation in R&D
5	Technological Entrepreneurship
6	Patent-Assertion Entities
7	Complex Approach

Because of the characteristics of universities, there are no patents transferred in authorized invention patents of Renmin University of China and Minzu University of China. Finally, 37 graphs are used to analyse the patent transfer modes. Mining the nature of each entity and relationships among the different entities through the network, seven kinds of technology transfer mode are summarized as shown in Table 2.

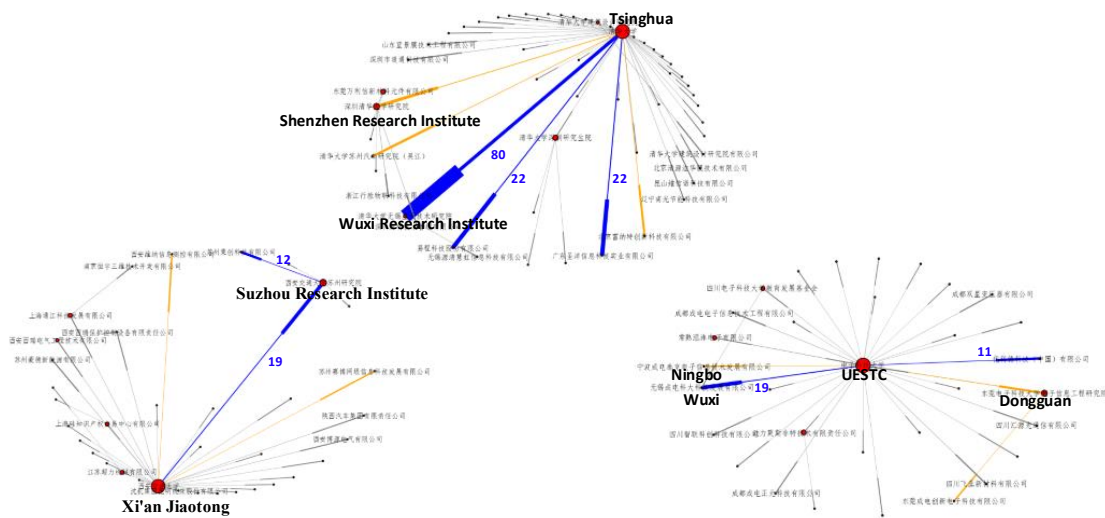


Figure 5. Examples for “Local Government-University Cooperation in R&D”

Tsinghua has built a series of research institutes with local governments based on local advantages, industrial concentration, market requirement and policy support. These institutes are public institutions with enterprise management and most patents in Tsinghua are transferred to these institutes. Xi'an Jiaotong University hatches a few enterprises in its Suzhou research institute with local government based on the technology support of different teams in school. UESTC builds public innovation platform with local governments (e.g., Wuxi, Dongguan and Ningbo) and that makes it develop fast in integration of resources, transformation of technological achievements and cultivation of talents.

CONCLUSION AND DISCUSSION

We classify each university into different transfer modes and count the number for each mode. Most universities belong to more than one mode and we choose the main modes for each university as shown in Table 3. Technology transfer in Chinese universities mainly concentrate on mode 1, mode 2 and mode 3. While in China there is science and technology progress law, which is praised as Bayh Dole Act of Chinese version (Lianhong He & Chen Jican, 2013). However, the law didn't make its proper effects, considering the period from issue to implementation. Patents, as professor's scientific research achievements, the ownership in China belongs to service invention. That's to say, the professor's patents belong to the professor's colleges and universities. While patents of universities belong to intangible assets, though their ownership are possessed, they are still in restraint of China state-owned assets management system, so it is really difficult to dispose separately. That is the reason why most patents of Dalian University of Technology, as a typical example of mode 2, have gone through a complex process in the transfer from school to enterprise and UTTOs play a very important role in mode 2.

Universities in mode 6 are relatively few, SIPO in China has issued two batches experimental organizations for patent operation, 70 in total, in order to improve the condition of Chinese patent management. Although more and more senior professors lead their teams to start a business, the basic conditions and atmosphere have not been established in Chinese universities, especially for excellent young teachers. Although core technologies are grasped in their hands, but insufficient funds seriously affect the development of innovation. The policies on talent development and appropriate legal environment are also the important factors that affect the success of entrepreneurship. However, with the deepening reform of China's S&T system, distribution of S&T resources will be more optimized. In the meantime,

technological entrepreneurship will become more feasible and common in the wave of “popular entrepreneurship and innovation”.

Table 3. Technology transfer modes of 985-project universities

<i>Transfer Mode</i>	<i>Counts</i>	<i>Universities</i>
1. Direct Transfer of Patents	10	BIT, BNU, SEU, ECNU, TJU, Shandong, Chongqing, NPU, Tongji, Jilin
2. Asset Management Unit of University	16	DUT, NEU, Fudan, HIT, SCUT, NJU, Xiamen(XMU), Tongji, Xi'an Jiaotong, USTC, CAU, Ocean University of China, SUN YAT-SEN UNIV., Chongqing, Jilin, HUST
3. University-Enterprise Cooperation in R&D	12	BUAA, UESTC, NKU, Tsinghua, Xiamen(XMU), Shandong, SJTU, TJU, NPU Zhejiang, CAU, NUDT
4. Local Government-University Cooperation in R&D	3	UESTC, Tsinghua, Xi'an Jiaotong
5. Technological Entrepreneurship	9	BUAA, Fudan, HIT, SJTU, Hunan, Sichuan, Zhejiang, NW A&F Univ, SUN YAT-SEN UNIV.
6. Patent-Assertion Entities	2	Lanzhou, Wuhan
7. Complex Approach	2	Peking, Central South Univ

REFERENCES

- Abreu, M., & Grinevich, V. (2013). The nature of academic entrepreneurship in the UK: Widening the focus on entrepreneurial activities. *Research Policy*, 42(2), 408-422.
- Rasmussen, E. (2008). Government instruments to support the commercialization of university research: Lessons from Canada. *Technovation*, 28(8), 506-517.
- Kalar, B., & Antoncic, B. (2015). The entrepreneurial university, academic activities and technology and knowledge transfer in four European countries. *Technovation*, 36, 1-11.
- Bekkers, R., & Freitas, I. M. B. (2008). Analysing knowledge transfer channels between universities and industry: To what degree do sectors also matter?. *Research policy*, 37(10), 1837-1853.
- Landry, R., Saihi, M., Amara, N., & Ouimet, M. (2010). Evidence on how academics manage their portfolio of knowledge transfer activities. *Research Policy*, 39(10), 1387-1403.
- D'Este, P., & Patel, P. (2007). University–industry linkages in the UK: What are the factors underlying the variety of interactions with industry? *Research Policy*, 36(9), 1295-1313.
- Grimaldi, R., Kenney, M., Siegel, D. S., & Wright, M. (2011). 30 years after Bayh–Dole: Reassessing academic entrepreneurship. *Research Policy*, 40(8), 1045-1057.
- Rao K, Meng X F, & Piccaluga A, et al. (2011). University Patent Technology Transfer in U.K. and its Referential Experiences. *Forum on Science and Technology in China*, (2), 148-154.

- Harmon, B., Ardishvili, A., Cardozo, R., Elder, T., Leuthold, J., Parshall, & Smith, D. (1997). Mapping the university technology transfer process. *Journal of Business venturing*, 12(6), 423-434.
- Markman, G. D., Phan, P. H., Balkin, D. B., & Gianiodis, P. T. (2005). Entrepreneurship and university-based technology transfer. *Journal of Business Venturing*, 20(2), 241-263.
- Grimpe, C., & Fier, H. (2010). Informal university technology transfer: a comparison between the United States and Germany. *The Journal of Technology Transfer*, 35(6), 637-650.
- Baldini, N., Grimaldi, R., & Sobrero, M. (2006). Institutional changes and the commercialization of academic knowledge: A study of Italian universities' patenting activities between 1965 and 2002. *Research policy*, 35(4), 518-532.
- Swamidass, P. M., & Vulasa, V. (2009). Why university inventions rarely produce income? Bottlenecks in university technology transfer. *The Journal of technology transfer*, 34(4), 343-363.
- Huang, W. L., Feeney, M. K., & Welch, E. W. (2011). Organizational and individual determinants of patent production of academic scientists and engineers in the United States. *Science and Public Policy*, 38(6), 463-479.
- Grimm, H. M., & Jaenicke, J. (2012). What drives patenting and commercialisation activity at East German universities? The role of new public policy, institutional environment and individual prior knowledge. *The Journal of Technology Transfer*, 37(4), 454-477.
- Zhang, X. Q. (2013). Four keys to implement the development strategy of innovation. Retrieved from http://www.china.com.cn/v/2014jjnh/2013-12/21/content_30963029.htm. Accessed 11 Jan 2015.
- Luan, C., Zhou, C., & Liu, A. (2010). Patent strategy in Chinese universities: a comparative perspective. *Scientometrics*, 84(1), 53-63.
- Song D L, Fu B H, & Tang H. (2011). Analysis of Patent Technology Transfer Modes of Univesities. *Forum on Science and Technology in China*, (3), 95-100.
- Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions and reversals. In *Soviet Physics Doklady*, 10(10), 707.
- Link, A. N., Siegel, D. S., & Bozeman, B. (2007). An empirical analysis of the propensity of academics to engage in informal university technology transfer. *Industrial and corporate change*, 16(4), 641-655.
- Chien, Colleen V. (December 10, 2012). Patent Assertion Entities. Presentation to the Dec 10, 2012 DOJ/FTC Hearing on PAEs. Available at SSRN: <http://ssrn.com/abstract=2187314>
- Lianhong He & Jican Chen (2013). The failure of the Bayh Dole Act of Chinese version and the ways of industrialization of intellectual property in chinese universities. *Intellectual Property (China)*, 27(3), 84-88.
-



Large Scale Disambiguation of Scientific References in Patent Databases

Kangran Zhao*, Emiel Caron** and Stanisław Guner**

*kangranz@outlook.com

School of Finance, Shanghai University of Finance and Economics, 777 Guoding Road, Yangpu District, Shanghai (China)

**caron@ese.eur.nl; guner.stan@gmail.com

Department of Econometrics, Erasmus School of Economics, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, (The Netherlands)

ABSTRACT

The PATSTAT database stores information on patent applications and publications. One of its tables, stores scientific references cited by patents. As such, this is a potentially powerful resource to investigate the relation between science, technology and innovation. We aim to provide a reliable way to conduct research on such databases. To this end, we employ automated data cleaning and extract bibliographic information. Furthermore, a scoring system is used, and clusters of duplicates of scientific references are obtained by a clustering algorithm.

INTRODUCTION

Two important questions arise in the analysis of scientific references cited by patents:

- How to obtain structured bibliographic information from poorly organized references?
- How to disambiguate scientific references?

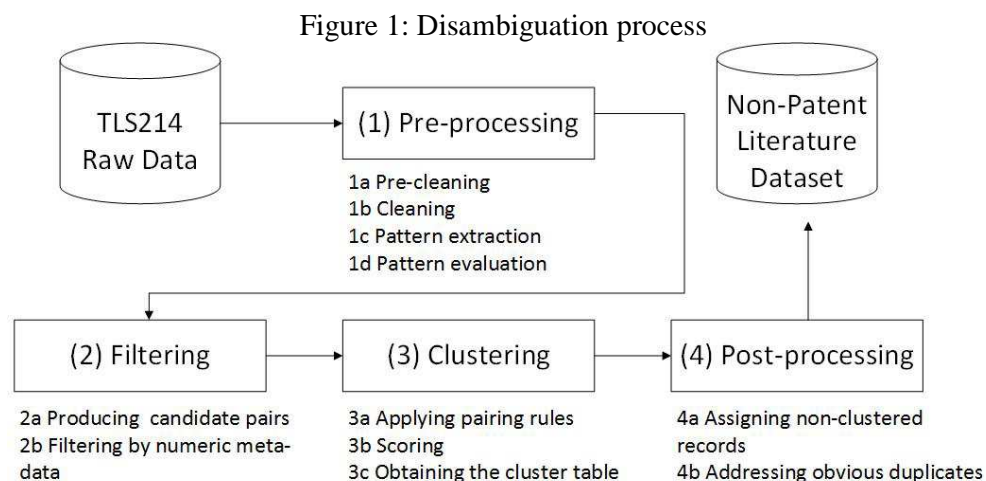
The answers facilitate research on scientometrics that provides measures for evaluation of scientific output through analysis of bibliographic information (Leydesdorff & Milojević, 2015). Unfortunately, such disambiguation systems for patent databases are not fully developed.

In this paper, we perform automated cleaning and then propose a disambiguation method using rule-based scoring and clustering in PATSTAT¹. The results of this method can facilitate policy evaluation and research on innovation, by providing views of connections of scientific references cited by patents. In the next section, we illustrate our methodology. After that we evaluate results with a golden set. We close the paper with some concluding remarks.

METHODOLOGY

Figure 1 illustrates an overview of our method, which aims to assign scientific records in a dataset to a cluster describing the same bibliography. The method is inspired by the work of Caron and van Eck (2014). We work with table ‘TLS214’ in the PATSTAT database. The number of records in TLS214 is 23,806,543 (version autumn 2014). In principle, operations performed on proper samples can be applied to the whole dataset. The number of records in the sample is 100,000. The project is conducted with Microsoft SQL Server. The code base is written in T-SQL in combination with C#.

¹ PATSTAT is a product of the European Patent Office. It is a periodical snapshot of patent related information organized in a relational database model.



Pre-processing

There are 16,853,440 unique records in the whole dataset, 69.5% of which are identified as scientific references. In step 1c, we mainly use regular expressions to extract information, based on labels, formats and ordering. Label based models use literal characters to narrow down the search. Format based models search for formats in which it is conceived to be written and one primary example is Format Based Names (Constans, 2009). Ordering based models extract information from where it is conceived to be written. Table 1 shows some of the aforementioned techniques for the record: “Harwood, C.F., Compaction Effect on Flow Property Indexes for Powders, J. Pharm. Sci., 60:161-163 (1971)”. In step 1d, we harmonize the extracted information.

Table 1. Techniques to parse the example record.

Type	Attribute	Technique	Extraction
Source	Journal Name	Format	J. Pharm. Sci.
		Ordering	
	Volume	Format	60
Author	Author Name	Format	Harwood, C.F.
		Ordering	
Title	Title	Ordering	Compaction Effect on Flow Property Indexes for Powders

Filtering

Most randomly matched pairs seem irrelevant. In step 2a, we obtain candidate pairs likely to be duplicates, if they meet any rule in Table 2.

Table 2. Filtering rules

Category	Attribute	Meta-data (used for filtering and pairing)
numeric	(every) numeric attribute (alone)	numbers (partial)
		sum of numbers and count of numbers
author	author name	author name and initials
alphabetic	alphabetic characters in bibliography	alphabetic characters (partial)
		alphabetic subtext at the start or the end
		the first or last word in alphabetic texts
title	title	title (partial)
		the first or last word in titles
source title	source title	source title (partial)
		the first or last word in source titles

Rule-based scoring

In step 3a, we measure similarity of record pairs in every attribute with the rules in Table 3. Numeric rules search for exact matches, while alphabetic rules adopt Levenshtein Distance (LD) or Jaccard Index (JI) algorithm with certain thresholds indicating strong string similarity. In step 3b, overall similarity is quantified as: Overall Score = Base points + Composite rules points - Negative rules points.

Table 3. Atomic rules

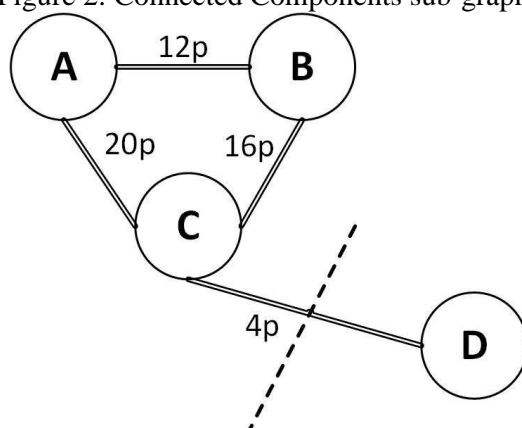
Numeric Rules				Alphabetic Rules	
No.	Attribute(s)	No.	Attribute(s)	No.	Attribute(s)
N1	d_year	N8	pages_start	W1	LD(name)
			pages_end		
N2	d_month	N9	xp_number	W2	LD(alphabetic)
N3	d_day	N10	issn	W3	LD(title)
N4	issue	N11	isbn	W4	LD(source_title)
N5	volume	N12	doi	W5	LD(Institution)
N6	edition	N13	accession	W6	LD(conference)
N7	url	N14	JI(bib_numeric_space)		

Compared to a fixed set of rules, a flexible subset for each pair can yield better model interpretability. For each pair, only three alphabetic rules with the highest points, as a subset, receive base points. For example, if one scores {W1:5, W2:-12, W3:5, W4:2, W5:1}, it receives $5 + 5 + 2 = 12$ points. A composite rule, in the following set, combines multiple atomic rules, e.g. {W: 1, 3, 4} \times {W: 1, 3, 4} \times {N: 4, 5, 6} + {W: 1, 3, 4} \times {N: 1} \times {N: 4, 5, 6, 8}. Composite rules are considered strong evidence. For instance, {W1, W3, N8} obtains pairs with the same page numbers, having similar author names and titles.

Clustering

We use an algorithm to find connected components to obtain clusters of pairs with scores above a certain threshold, see the example in Figure 2.

Figure 2: Connected Components sub-graphs



Only ABC cluster is above the 8 point threshold and D is not added due to lack of proof.

Post-processing

In step 4a, records for which no duplicates were detected, are assigned to new, single record clusters. Once the representative of a group of duplicates is assigned, we append the remaining duplicates from the pre-cleaning to the same cluster.

EVALUATION

We use precision and recall analysis to evaluate the clusters on a “Golden Set”², which is a verified sample of records in TLS214. The results of the analysis are depicted in Table 4, which shows precision and recall values for extracted attributes, and Table 5, which shows the number of duplicates per cluster in the sample. The total processing time for the sample was 24 minutes.

Table 4. Evaluated attribute vs. Golden set

Evaluated attribute vs. Golden set			Extraction rate	
Attribute	Precision	Recall	Our procedure	Golden Set
title	0.926	0.908	0.801	0.818
source_title	0.801	0.809	0.833	0.824
name	0.898	0.911	0.92	0.906
d_year	0.988	0.99	0.961	0.959
volume	0.919	0.891	0.67	0.691
issue	0.924	0.91	0.304	0.309
pages_start	0.956	0.933	0.818	0.837
pages_end	0.969	0.929	0.734	0.765

² We would like to thank Jos Winnink of the Patent Statistics Office of the Dutch Central Planning Bureau and the Centre for Science and Technology Studies for providing the golden set.

Table 5. Cluster statistics

Amount of duplicates per cluster	Amount of Clusters	%
42	1	0.03%
9	3	0.09%
16-19	5	0.16%
10	5	0.16%
8	15	0.47%
11-15	16	0.50%
7	17	0.54%
6	26	0.82%
5	46	1.45%
4	113	3.56%
3	410	12.92%
2	2516	79.29%
TOTAL of Multiple clusters	3173	100.00%
1	61086	

Table 6 presents the visual inspection of an example cluster, which shows name variants found in patents of the scientific reference of Köhler and Milstein (1975). Notice the differences in citation style.

CONCLUSION

In this paper we developed a disambiguation method for large patent databases. Prior to scoring, we performed automated cleaning, where we obtained bibliographic attributes, some of which were even in ambiguous patterns, such as titles. The filtering and scoring systems have a good coverage and the obtained clusters solved the problem of record disambiguation for a large representative sample. In future research, we want to apply the method to disambiguate the whole TLS214 table in PATSTAT.

Table 6: Example cluster.

cluster	new_id	npl_biblio
100	2503	KOHLER; MILSTEIN NATURE vol. 256, 1975, page 495
100	9107	KOWER ET AL. NATURE vol. 256, 1975, page 495
100	9672	Köhler and Milstein, Continuous cultures of fused cells secreting antibody of predefined specificity, Nature 256(5517):495-497, Aug. 1975.
100	10733	Kohler et al., Continuous cultures of fused cells secreting antibody of predefined specificity, Nature, 1975, 256:495.
100	11041	NATURE vol. 256, 1975, page 495
100	26554	Kohler, G. et al., Continuous Cultures of Fused Cells Secreting Antibody of Predefined Specificity, Nature, 256: 445-497 (1975).
100	40683	Nature 256:495 497, 1975, Continuous cultures of fused cells secreting antibody of predefined specificity, Kohler et al.
100	42384	KOHLER ET AL. NATURE vol. 256, 1975, page 495
100	49871	Kohler, G. et al., Continuous cultures of fused cells secreting antibody of predefined specificity, Nature, vol. 256, pp. 495-499, Aug. 7, 1975.
100	52968	Kohler, G. et al., Continuous cultures of fused cells secreting antibody of predefined specificity, Nature, Aug. 7, 1975, vol. 256, pp. 495-497.
100	66744	Köhler et al., Continuous cultures of fused cells secreting antibody of predefined specificity, Nature 256:495-497 (1975).
100	71727	KÖHLER G; MILSTEIN C: 'Continuous cultures of fused cells secreting antibody of predefined specificity' MATURE vol. 256T, 1975, pages 495 - 7
100	73829	G. Kohler and C. Milstein, Continuous Cultures of Fused Cells Secreting Antibody of Predefined Specificity, Nature 256:495 (1975).
100	76102	KOHLER; MILSTEIN NATURE vol. 256, 1975, pages 495 - 497
100	89063	Kohler and Milstein, Nature 256:495 (1975).
100	92306	Kohler and Milstein, Continuous cultures of fused cells secreting antibody of predefined specificity, Nature 256: 495-497 (1975).

REFERENCES

Caron, E., van Eck, N.J., (2014). Large scale author name disambiguation using rules-based scoring and clustering, In Proceedings of the 19th International Conference on Science and Technology Indicators, pages 79-86, Leiden, The Netherlands.

Constans, P. (2009). A Simple Extraction Procedure for Bibliographical Author Field. <http://arxiv.org/abs/0902.0755>.

Leydesdorff, L., Milojević, S. (2015). Scientometrics. In M. Lynch, International Encyclopedia of Social and Behavioral Sciences (Vol. Science and Technology Studies). Elsevier.

Köhler G., Milstein C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. Nature 256. Aug 7, 256, 495 – 497, doi:10.1038/256495a0.

POSTER SESSION 2

**Data Characterisation, Classification,
Visualisation and Indicator Design**



Quantifying and visualizing different types of scientific collaboration in Nanoscience and Nanotechnology field¹

Zaida Chinchilla-Rodríguez^{*}, Sandra Miguel^{**}, Antonio Perianes-Rodríguez^{***}, Carlos Olmeda-Gómez^{***} and María-Antonia Ovalle-Perandones^{***}

^{*} zaida.chinchilla@cchs.csic.es

Instituto de Políticas y Bienes Públicos. Consejo Superior de Investigaciones Científicas, C/ Albasanz, 26-28. 28037 Madrid (Spain), SCImago Research Group

^{**} smiguel@fahce.unlp.edu.ar

Instituto de Investigaciones en Humanidades y Ciencias Sociales—IdIHCS (UNLP-CONICET), Facultad de Humanidades y Ciencias de la Educación, Universidad Nacional de La Plata, calle 51 e/124 y 125, Ensenada, 1925, Argentina, SCImago Research Group.

^{***} antonio.perianes@uc3m.es, carlos.olmeda@uc3m.es, antonia.ovalle@uc3m.es

Departamento de Biblioteconomía y Documentación, Universidad Carlos III, C/ Madrid, 128. 28903 Getafe (Spain), SCImago Research Group

ABSTRACT

The general aim of this study is to quantify the benefit rate in visibility and impact of scientific production in the field of N&N bearing in mind the different types of output (total, in leadership, excellent, and excellent with leadership) of the six main producers of knowledge in N&N in Latin America in the period 2003-2013. More specifically we aspire to visualize the networks of international collaboration in a given country (ego-network) to represent the difference between the citations received per type of output, and identify the associates with whom a country has greater potential and capacity to generate knowledge of high quality, as well as the differences existing in terms of visibility depending on the type of production analyzed. In short, we wish to determine the benefits of such collaborative efforts. In this way we could respond to questions such as: a) With which countries is collaboration established? and b) With which collaborating countries are the greatest volume of citations per document obtained, according to the type of output.

INTRODUCTION

International collaboration (IC) in the creation of knowledge is responsible to change the structural stratification of science. This change in the relation between the geographical and intellectual dimensions of science has profound implications for the science (Leydesdorff et al. 2013).

Analysis of collaboration in Latin American countries is of particular significance, because initiatives are often the result of “research-for-aid” arrangements, generally based on North–South asymmetries (Bonfiglioli & Mari, 2000). Collaboration for mutual benefit and excellence has gained increasing acceptance, with “partner” selection becoming a strategic to enhance one’s own production (Velho, 2002).

¹ This work was made possible through financing by the Project NANOMETRICS (Ref. CSO2014-57770-R) supported by Ministerio de Economía y Competitividad of Spain.

Collaboration is an added value for increasing productivity and visibility (Gazni, Sugimoto & Didegah, 2012). It has also been demonstrated that countries benefit from participation in multinational projects, which ultimately leads to an improved citation factor (Glänzel & De Lange, 2002). Furthermore, increasing international collaborations has a positive effect on the impact factor and the research quality of publications (Wagner et al., 2001). The effects of collaboration don't always translate into benefits at the same magnitude.

Whatever type of collaboration may be quantified through collaborative projects, publications in common, informal contacts, the interchange of researchers or fellows among different countries (Guerrero, Olmeda & Moya, 2013).

OBJECTIVES

The general aim is to quantify the benefit rate in impact of scientific production in the field of nanoscience and nanotechnology (N&N) bearing in mind the different types of output (total, in leadership, excellent, and excellent with leadership) in Latin America (2003-2013). We aspire to visualize the heliocentric-networks of Argentina international collaboration to represent the difference between the citations received per type of output.

MATERIAL AND METHODS

The data set was obtained from SCImago Journal & Country Rank (<http://www.scimagojr.com/>) and SCImago Institutions Rankings (<http://www.scimagoir.com/>), based on the Scopus database. The indicators are:

- Output (ndoc).
- % Collaboration types (No-collaboration; National collaboration; International & National collaboration; International collaboration; International collaboration with leadership).
- % Leadership types.
- Cites per document: total (cpd) and leading papers (cpd_L).
- Normalized citation impact (NI)
- Normalized citation impact with leadership (NIL) (Moya et al 2013).
- Benefit rate of collaboration in normalized citation impact (BRCNI)

RESULTS

Latin America published 4,811 documents in the category N&N. Figure 1 represents 2.73% of the world output. Brazil is the country with the most output followed by Mexico and Argentina.

Figure 1. Scientific output in N&N in main six Latin American countries

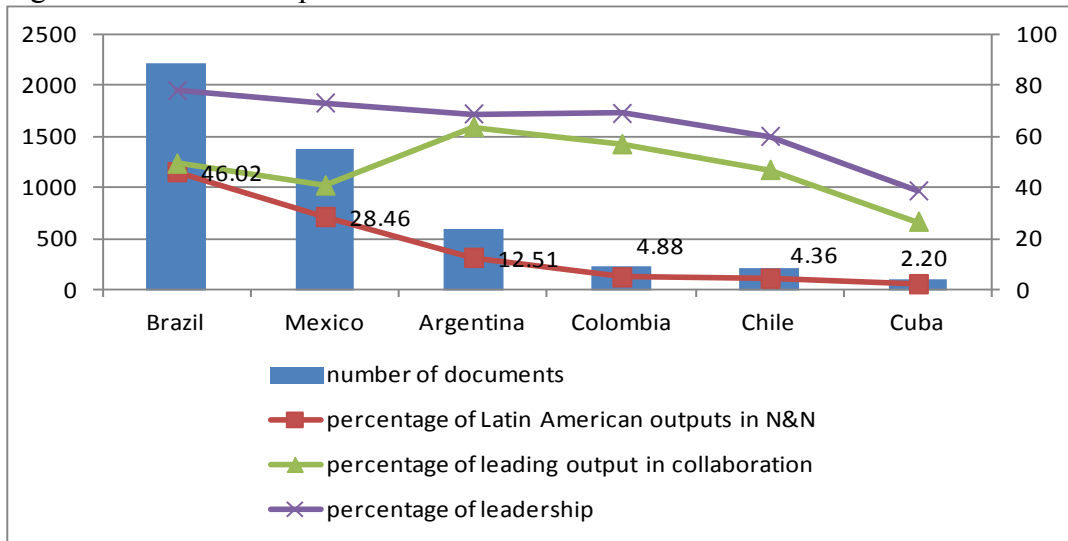
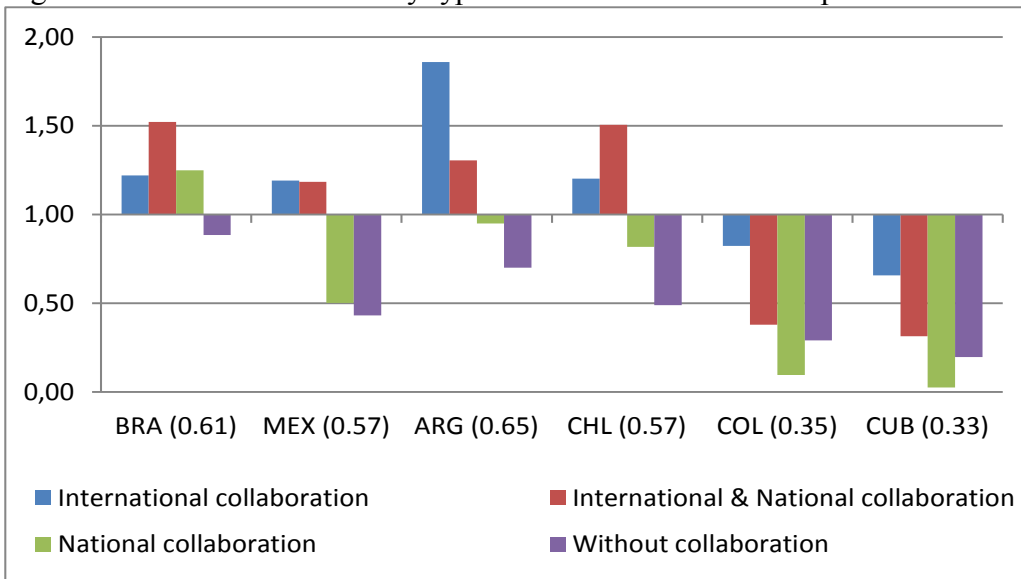


Figure 2 shows the normalized citation by type of collaborations of all outputs and figure 3 in leading outputs, the value of world mean normalized impact is 1. The number between parentheses after the country's abbreviation is the impact reached by each country.

Figure 2. Normalized citation by type of collaboration of all outputs



Brazil, Argentina and Chile are the most dependent countries upon international collaboration to obtain good results in terms of yield, and are therefore the ones that most benefit from this association.

Figure 3. Normalized citation by type of collaboration of leading outputs

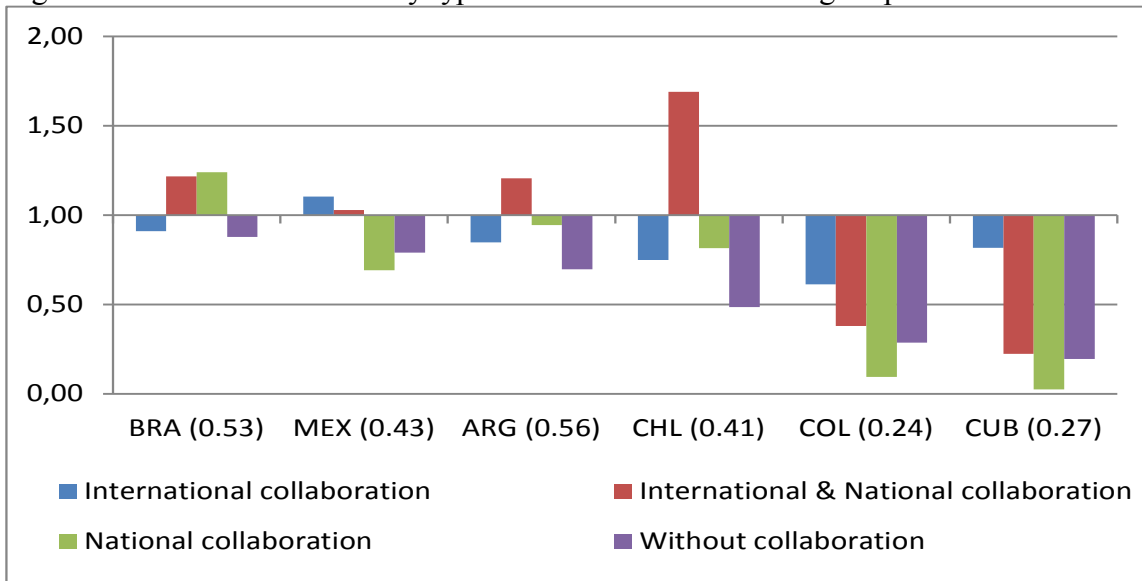
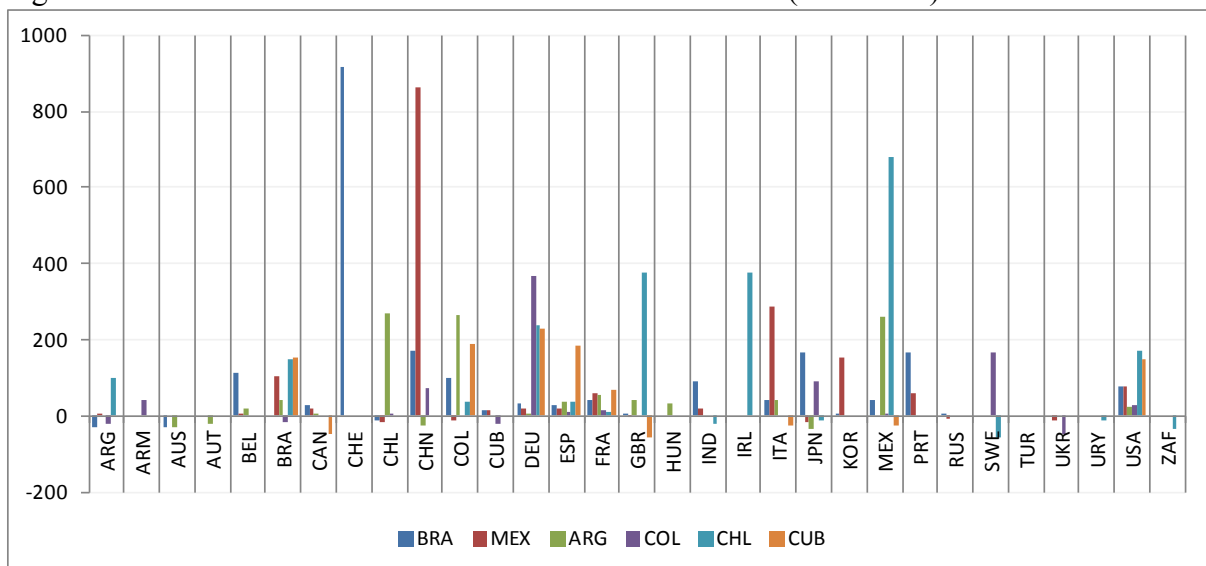


Figure 4 presents the foreign countries involved in collaboration and the index of benefit in normalized impact (BRCNI). Argentina is a good associate for Mexico, Chile, and Colombia, who benefit as well from the impact of the output undersigned by this country. The autonomy of Argentina in obtaining greater citation when it leads research with Austria, China and Japan is also remarkable.

Figure 4. Benefit rate of collaboration in normalized citation (NI vs NIL)



SCIENTIFIC IC AND CPD BY TYPES OF PRODUCTION IN ARGENTINA

The heliocentric representations depict the international collaboration of Argentinean scientific output in N&N. Thus, one can quickly spot with which countries more is published (greater volume) and with which one is more visible (closer to the center) (Chinchilla-Rodríguez, et. al., 2010).

In the four heliocentric networks (figures 5, 6, 7 and 8) the countries located into the red orbit are the more cited and put Argentinean research on the map of excellence achieving the highest visibility and international impact.

Figure 5. Network of IC in all output

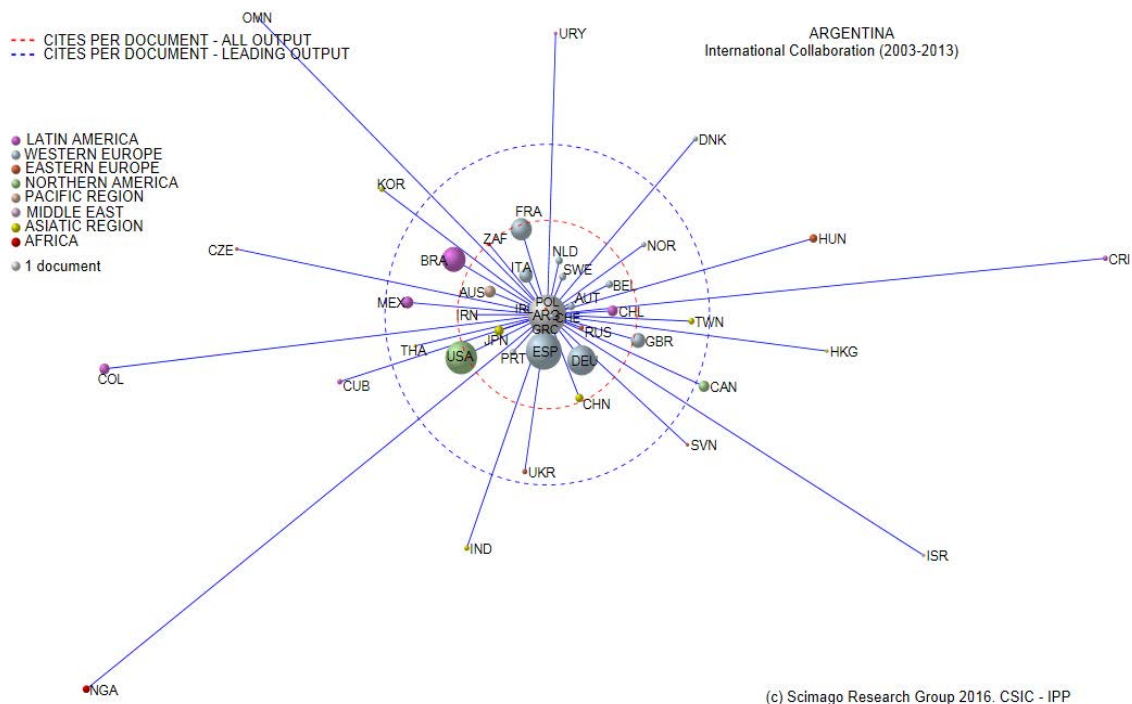


Figure 6. Network of IC in excellence output

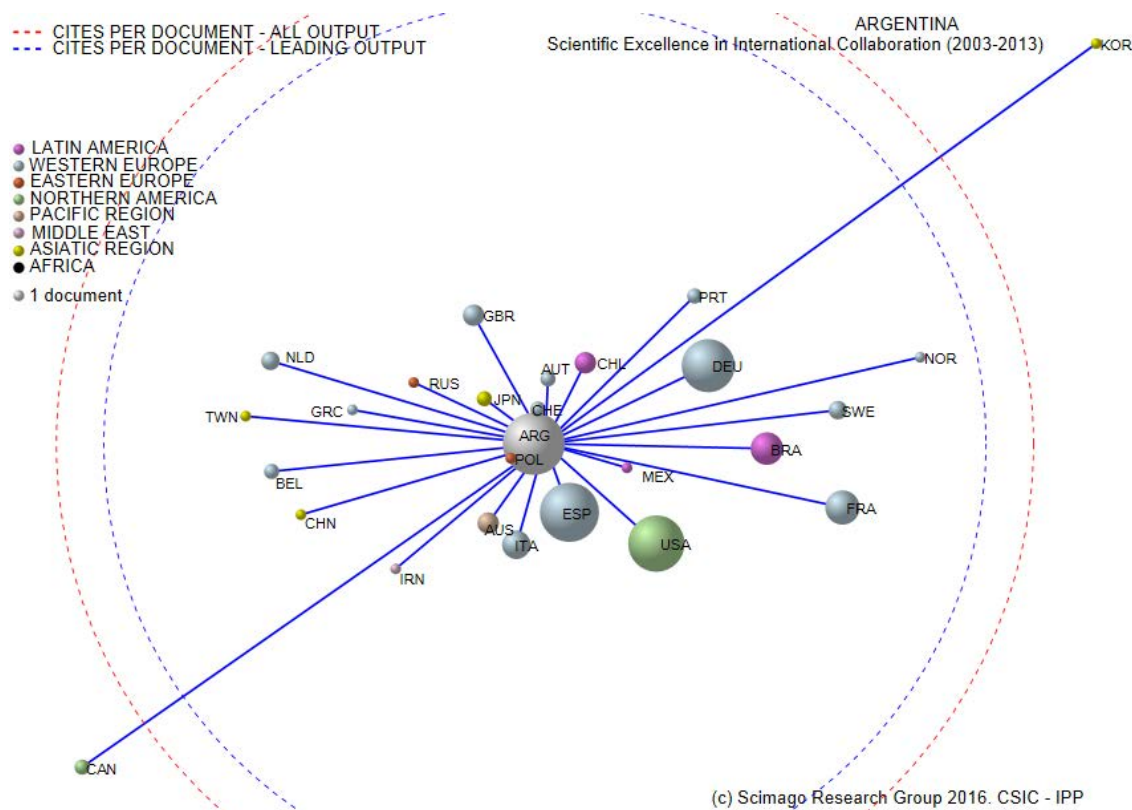


Figure 7. Network of IC in leading output

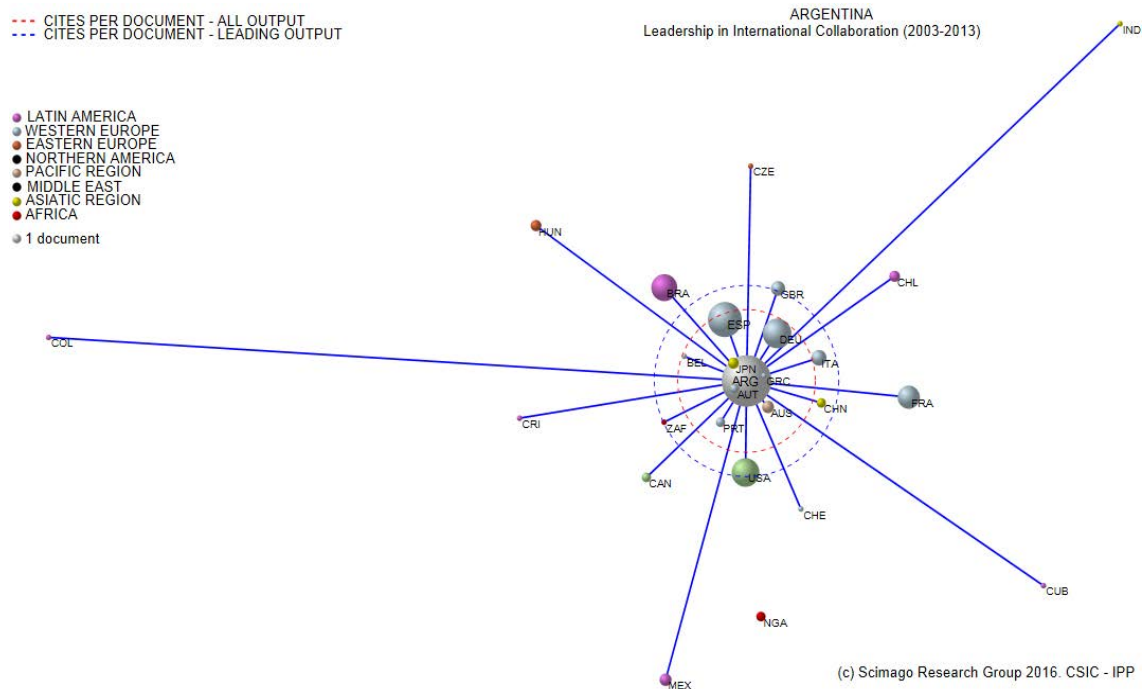
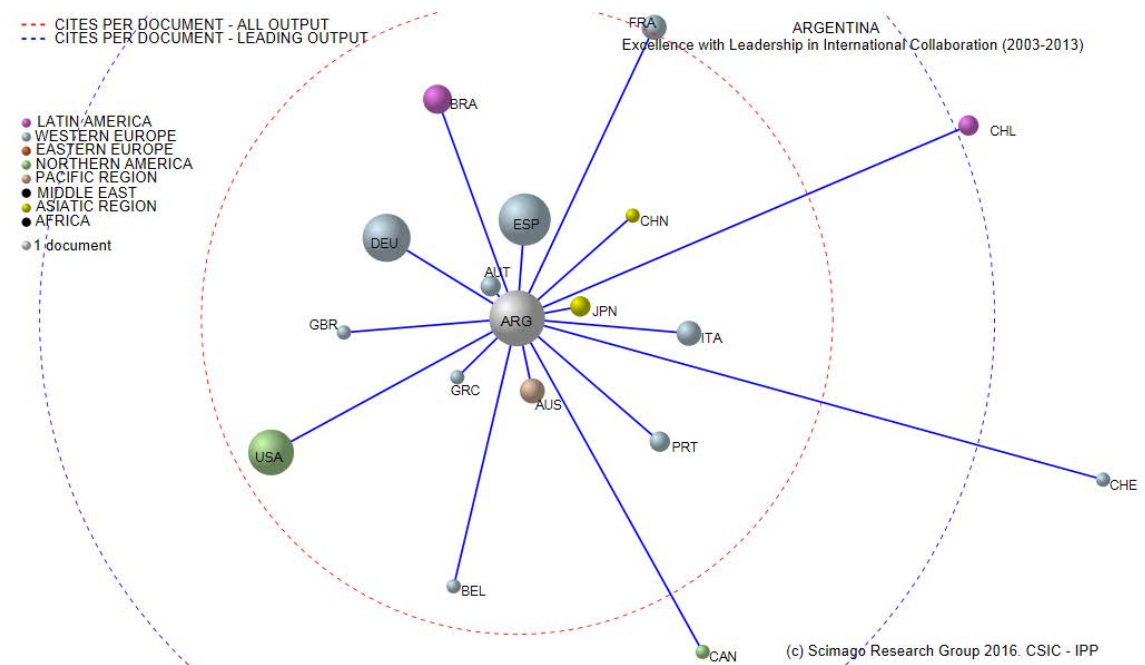


Figure 8. Network of IC in excellence with leadership output



REFERENCES

- Bonfiglioli, A. & Mari, E. (2000). La cooperación científico-tecnológica entre la Unión Europea y América Latina: el actual contexto internacional y el Programa Marco de la Unión Europea. *Revista REDES*, 7, 183–208.
- Chinchilla-Rodríguez, Z., Vargas-Quesada, B., Hassan-Montero, Y., González-Molina, A., & Moya-Anegón, F. (2010). New Approach to the Visualization of International Scientific Collaboration. *Information Visualization*, 9, 277-287.
- Gazni, A., Sugimoto, C.R., Didegah, F. (2012). Mapping World Scientific Collaboration: Authors, Institutions, and Countries. *Journal of the American Society for Information Science and Technology*, 63(2): 323–335.
- Glänzel, W. & De Lange, C. (2002). A distributional approach to multinationality measures of international scientific collaboration. *Scientometrics*, 54(1): 75-89.
- Guerrero-Bote, V.P., Olmeda-Gómez, C., Moya-Anegón, F. (2013). Quantifying the benefits of international scientific collaboration. *Journal of the American Society for Information Science and Technology*, 64 (2):392-404.
- Leydesdorff, Loet; Wagner, Caroline S.; Park, Han-Woo; Adams, Jonathan (2013). International collaboration in science: the global map and the network. *El profesional de la información*, 22, 87-94.
- Moya-Anegón, F., V. Guerrero-Bote, L. Bornmann, and H.F. Moed. (2013). “The Research Guarantors of Scientific Papers and the Output Counting: A Promising New Approach.” *Scientometrics*, 97(2): 421–34.
- Velho, L. (2002). North-south collaboration and systems of innovation. *The International Journal of Technology Management and Sustainable Development*, 1, 25–49.



Internal Migration of Scientists in Russia and the USA: the Case of Physicists¹

Ekaterina L. Dyachenko*

* edyachenko@hse.ru

Institute for Statistical Studies and Economics of Knowledge, National Research University Higher School of Economics, Myasnitskaya st., 9/11, office 443, Moscow, Russia

ABSTRACT

There is a shared understanding that migration flows influence a country's scientific performance. Although there are plenty of studies on academic mobility, most of them focus on international migration, while inter-regional and inter-organizational mobility seem to be unduly neglected. In this study we compare the structure of internal labor mobility for Russian and American scholars working in the field of applied physics. The goals are to describe the networks of mobility of Russian and American physicists and to formulate hypotheses of how the features of network are connected to scientific performance.

INTRODUCTION

The constant growth of the mobility of scientists is one of the major trends in how science is made today. This growth is assumed to be one of the factors of scientific and technological progress. There is a shared understanding that academic/ inventors mobility influences total scientific performance/ innovation [Kaiser et al., 2015]. Researchers and policy makers are interested in learning the exact patterns of that influence, the structure of mobility flows, and the predictors of mobility scale and directions. There are plenty of studies on academic mobility, most of which focus on international migration, while within-the-borders mobility seem to be unduly neglected. In this study we compare the structure of within-the-borders labor mobility for Russian and American physicists.

When scientists change jobs they bring to their new workplace the experience, tacit knowledge and social ties acquired at the previous workplace [Almeida & Kogut, 1999; Agrawal et al., 2006, Trippl, 2013]. It is not only an organization which receives new knowledge when a new employee joins it. Migration of a scientist to new city or country generally enriches these places beyond the particular institution [Almeida & Kogut, 1999; Oettl & Agrawal, 2008; Trippl, 2013]. Apart from this spillover effect there is an effect of backward knowledge flow which means that research community can benefit not only from those who join it, but also from those who leave it [Agrawal et al., 2006]. With this assumption that migration of a researcher creates complex multidirectional knowledge flow we can consider the national science as the network of cities connected by the knowledge flows induced by the researchers who migrate between cities.

We suppose that the mobility-induced knowledge transfer is a significant factor of the overall performance of the knowledge production system. It is not only the intensity of labor mobility

¹ The work was prepared within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE) and the subsidy granted to HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program.

which is important, but also the topology of mobility network. For example, knowledge transfer should occur differently in star-like networks and in decentralized networks, and this difference probably matters for the total outcome. Generally, one can speak of relatively more or relatively less productive labor mobility. In this study we use Social Network Analysis approach to analyze migration of researchers, specifically of scholars working in the field of applied physics. The goals are to describe and compare the networks of mobility of Russian and American physicists and to formulate hypotheses of how the features of network are connected to national scientific performance.

Both Russia and the USA are major players in the field of applied physics research, with thousands of researchers and large networks of organizations performing R&D in this field. In both countries research institutions are widely dispersed geographically. Apart from these similarities, there are many differences in how science is organized in the two countries as well as in internal migration trends. The USA is known as one of the most mobile countries while Russia is characterized by a low level of academic mobility. In this study we are interested not so much in estimating the difference of mobility levels in two countries but more in comparing the structure of internal mobility.

METHODOLOGY

To obtain the data on researchers' affiliations we used academic papers from applied physics field, published in 2009 and indexed in the Web of Science database. From these papers we randomly selected 200 papers with authors affiliated with Russian institutions and 200 papers with authors affiliated with American institutions. Thus we obtained the set of 594 authors who worked in Russia in 2009 and 674 authors who worked in the USA in 2009. For each researcher we obtained the data on his/her affiliations from the papers published in 2013-2014. The data were available for 424 Russian authors and 396 American authors. To distinguish between researchers with the same name we used CVs published on the web, institutional websites and Russian Science Citation Index. The affiliations of the researchers in 2009 and in 2013-2014 were used to build the networks of mobility, where the nodes represent the cities and the edges represent the researchers who migrated between 2009 and 2014. The networks were visualized and analyzed by means of UCINET software. This study does not set the task to establish some facts on the direction or scale of the internal mobility flows between particular cities in Russia and the USA. That would demand much larger samples of researchers. What we are interested in here is the difference in topology of networks, which can be associated with the overall performance of the system.

RESULTS

Russia and the USA differ significantly in how authors are distributed across cities. In Russia 594 authors of 2009 papers are dispersed across 43 cities, in the USA 674 authors of 2009 papers are distributed across 138 cities. Such a difference stems in the first place from the difference in size of the Russian and American applied physics research sectors – in the USA the sector is larger in terms of the number of researchers and research centers. When we compare two distributions of authors across cities, we see that the two sectors differ not only in size. The distribution of Russian authors is highly skewed, with Moscow accounting for 27% of the sample, and three leading cities (Moscow, St. Petersburg, Nizhnii Novgorod)² accounting for more than half. American authors are distributed more evenly. There is no city which is home to more than 5% of the authors.

² We use Web of Science spelling.

Figure 1 and Figure 2 show the distribution of Russian and American physicists across cities. The size of the circle reflects the number of researchers working in the city in 2009. Compared to the USA in Russia the cities where physicists work are scattered in a more uneven way. There is concentration of such cities in Central Federal District of Russia, especially around Moscow, which is the biggest point on the map. In the USA, the cities where physicists work are distributed more evenly, there are several clusters in the different regions of the country with the North-East as high-density area.

Figure 1. Geographical distribution of Russian physicists from the sample, 2009.

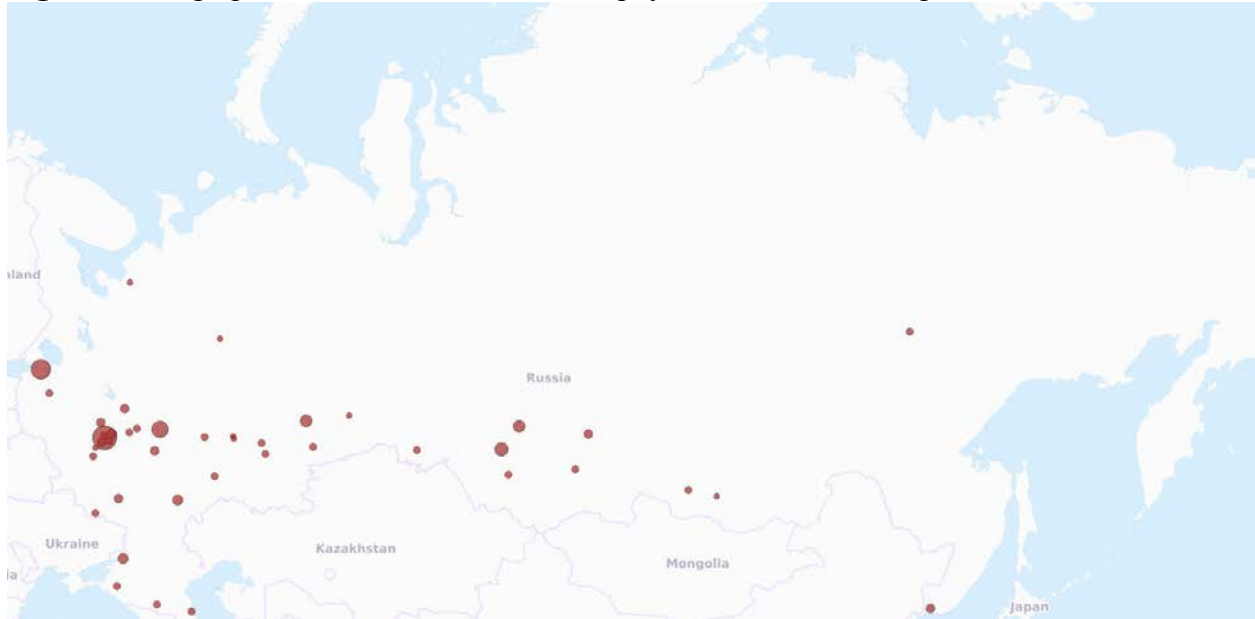
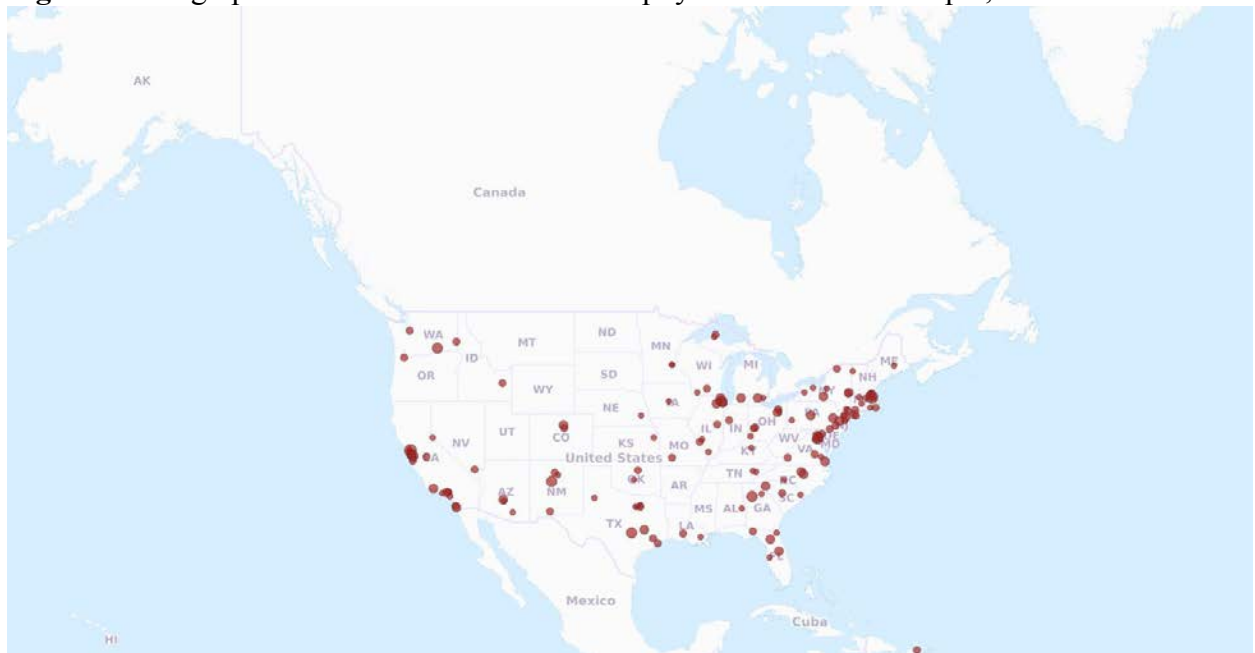


Figure 2. Geographical distribution of American physicists from the sample, 2009.



The figures in the Table 1 point at important differences in academic mobility of two countries, both quantitative and qualitative. The first important thing is that the mobility rates are more than two times higher in the USA than in Russia, which is true for the overall mobility, for mobility inside the country, and for international mobility. The qualitative difference between mobility patterns in two countries is that for American researchers the internal mobility is mostly ‘complete’, while for Russian researchers it is mostly ‘partial’³. Researchers with multiple affiliations represent another mode of labor mobility. While shuttling between cities and collaborating with the colleagues in these cities they also establish the knowledge flows. Our data show that these two modes of mobility - job-changing mobility and job-combining mobility – are related. For both, Russia and the USA, the share of those who combine jobs in several cities is much higher among mobile researchers than among immobile. While understanding that beyond quantitative indicators of mobility there can be a wide variety of practices, which shape the knowledge flows very differently, still we can formulate in quantitative terms the general hypothesis that performance of the knowledge production system is positively related to its labor mobility level. At the same time, we can expect that the relationship is more complicated than “the more mobility the better”.

Table 1. Mobility rates in the samples of Russian and American physicists.

	Russia	USA
Authors in the sample	594	674
Authors with full data on affiliations (2009 and 2013-2014)	424	396
Authors who changed cities between 2009 and 2014 (%)	49 (12%)	123 (31%)
Authors who went internationally mobile between 2009 and 2014 (%)	12 (3%)	39 (10%)
Authors who changed cities inside the country (%)	34 (8%)	76 (19%)
Authors who completely changed cities inside the country (%)	16 (4%)	64 (16%)

* % of the number of authors with full data on affiliations (424 and 396 respectively).

Figure 3 and Figure 4 show the networks of internal incoming mobility for Russia and the USA. Nodes towards which the edges are directed are the cities where people came to. For our samples of researchers which are of nearly the same size for Russia and the USA we see that there are five times more cities in the USA involved in mobility network than in Russia. In case of Russia the network has a center-periphery structure, where Moscow and Tomsk have diverse incoming flows of researchers while other cities either do not accept migrants at all or have sparse incoming migration flows. The network for the USA consists of many unconnected components. Like in case of Russia in the USA only a few cities are characterized by diverse incoming migration (Berkeley, Cambridge, Argonne). Still, the skewness of distribution of migrants across cities is much higher for Russia than for the USA. In Russia, mobile researchers generally go to Moscow. Another city with relatively diverse incoming migration, Tomsk, is probably only temporarily attractive for migrants⁴. In the USA mobile researchers go “everywhere”, there is no general direction of migration.

³ Many scientists in our samples have multiple affiliations, and we registered any change in an author’s city affiliations.

⁴ The physicists from our sample who moved to Tomsk between 2009 and 2014 got their new jobs at the universities, which participate in programs of the development of major universities run by Russian government since 2008. One can doubt whether the trend of migrating to Tomsk outlasts these particular programs.

Figure 3. Network of internal incoming mobility of Russian physicists.

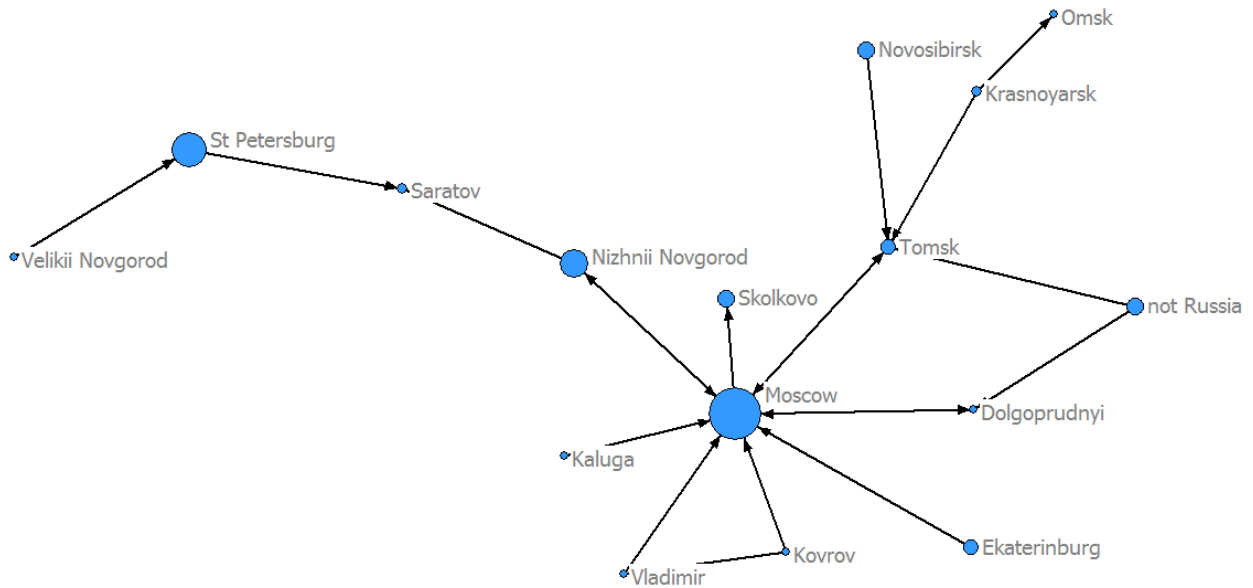
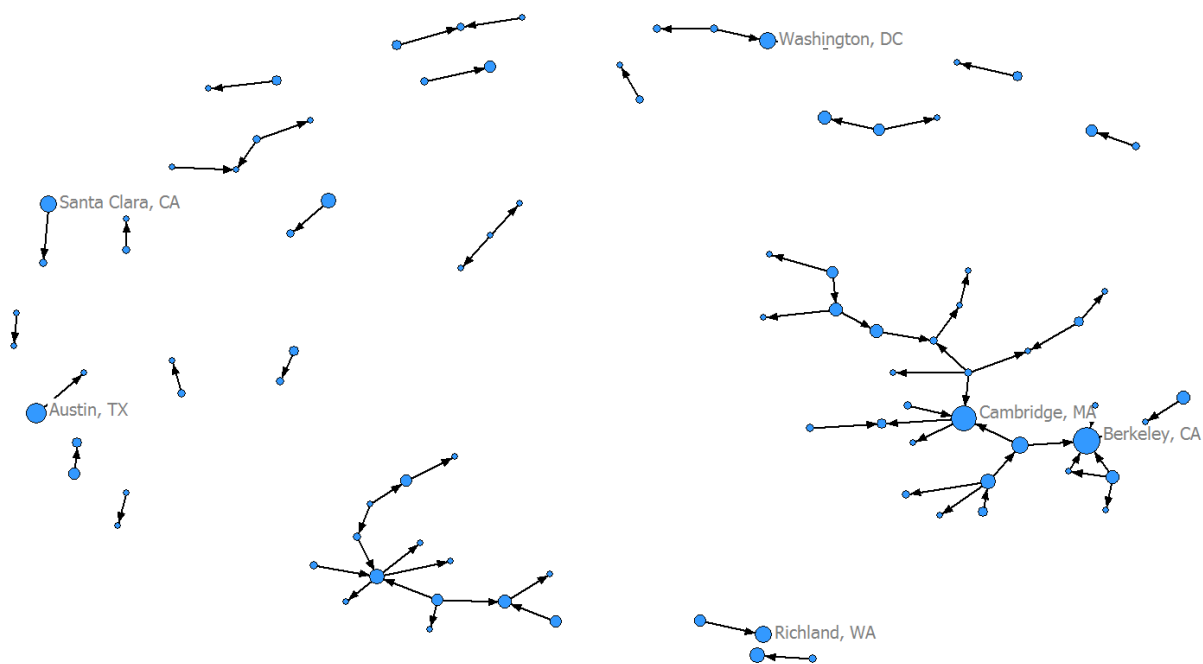


Figure 4. Network of internal incoming mobility of American physicists.



The skewness of internal mobility distribution can be important when we consider how the network structure influences the performance of the whole system. While it is possible that a migrant researcher can generate the flow of knowledge **to** the city he had **left**, still it seems plausible to consider the cities which have a sparse incoming flow of researchers as disadvantaged. We can expect that research community in such a city consists almost entirely of locals, hence it has weaker ties with the outer world than a community of researchers with geographically diverse education and work experience [Scellato et al., 2015]. In a knowledge production system where most communities are quasi-isolated, the lack of integration can cause the lack of efficiency.

The flows of migrants should depend on how the research institutions are distributed across cities. It seems natural – and not dysfunctional – for big cities (those with many jobs for researchers) to attract more migrants, than cities with few research institutions attract. The perfectly uniform distribution of mobility probably is not the most productive unless the distribution of research centers across cities is also uniform, which is never the case. What is important is whether big cities attract more than they should have according to their size. If they do, it leaves the periphery underfed with external labor force, which lowers the efficiency of the whole knowledge production system.

When we look at the positions of the biggest cities in mobility networks for Russia and the USA, we see that in both countries the largest cities – Moscow and Berkeley – have the largest amount of incoming migration. But also we can see the difference between the two countries considering the role of the big cities in the migrations networks. In Russia, Moscow attracts 45% of migrants while hosting 27% of physicists from the sample. Top 10% of the biggest cities in Russia attract more than half of all migrants, while in the USA, the similar proportion of the biggest cities accounts for only 22% of migrants. The whole distribution of migrants in Russia shows that not only the majority of cities lack incoming migration but this deficit concerns in the first place those already disadvantaged – cities with a modest research community tend to receive little “new blood”.

CONCLUSION

In this study our interest is driven by the widely-discussed idea that labor migration is one of the mechanisms of knowledge transfer. We assume that neither the rapid development of telecommunication technologies nor the growing ease of travelling do not diminish the role of labor mobility in knowledge transfer. Hence we expect the structure of internal migration to be an important factor affecting the performance in knowledge-intensive sectors. We found that researchers working in the field of applied physics in the USA are characterized by a higher level of mobility and more equally distributed mobility than physicists in Russia. We suppose that both the intensity level and the diffusion of mobility flows are positively related to the performance of the sector. In order to confirm or reject these assumptions, as well as to estimate the scale of the effect, further research is needed.

We have already mentioned that internal academic mobility lacks attention of researchers compared to international mobility. In Russia this disparity is echoed in the agenda of science policy. The internal mobility of researchers, low as it is, remains at the margins of research and policy-making. We chose the USA as a model for comparison with Russia not only because it is the R&D leader but because it is characterized by a high level of academic mobility. The point we want to make in this study is that not only the overall level of mobility matters. It is also important how mobility flows are distributed. The structure of labor mobility network in Russia makes us think that if a national science has a center-periphery structure, there is a risk for the periphery being trapped into becoming no more than a resource base for the centers. In such a case, the labor migration does not serve as a knowledge-exchange mechanism, but rather as a mechanism for the reproduction of inequality. While the big centers attract relatively diverse and intense flows of incoming migrants, few mobile researchers choose to move to the peripheral cities. The institutions in such cities are forced to employ on the local market, which means in the case of knowledge production that they remain semi-isolated and in the end disadvantaged, which stimulates researchers to leave them for the big centers.

Instruments for network analysis are not widely used in studies of scientific migrations, especially when within-borders migration is in the focus. We are not aware of any studies in which knowledge transfer is analyzed through the lens of mobility networks in order to define how network structure relates to national scientific performance. The study presented here, although it does not establish causal relations, produces some insights from the comparison of two networks and sets an agenda for future research.

REFERENCES

- Agrawal, A., Cockburn, I., & McHale, J. (2006). Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships. *Journal of Economic Geography*, 6(5), 571-591.
- Almeida, P., & Kogut, B. (1999). Localization of knowledge and the mobility of engineers in regional networks. *Management science*, 45(7), 905-917.
- Kaiser, U., Kongsted, H. C., & Rønde, T. (2015). Does the mobility of R&D labor increase innovation? *Journal of Economic Behavior & Organization*, 110, 91-105.
- Oettl, A., & Agrawal, A. (2008). International labor mobility and knowledge flow externalities. *Journal of International Business Studies*, 39(8), 1242-1260.
- Scellato, G., Franzoni, C., & Stephan, P. (2015). Migrant scientists and international networks. *Research Policy*, 44(1), 108-120.
- Trippl, M. (2013). Scientific mobility and knowledge transfer at the interregional and intraregional level. *Regional Studies*, 47(10), 1653-1667.



The Global Research Identifier Database GRID – Persistent IDs for the World’s Research Organisations

Martin Szomszor^{*}, Andres Mori

^{*} m.szomszor@digital-science.com; a.mori@digital-science.com

Digital Science, London, N1 9FN, UK

INTRODUCTION

The Global Research Identifier Database (GRID) <https://grid.ac> is a free (CC-BY), manually curated catalogue of research organisations designed to support disambiguation, integration and analysis of data associated with activities featuring in the scientific process. Central to this effort is the allocation and publication of *persistent identifiers* to support unambiguous attribution even when names, locations and organisational structures may change over time. The need for such a reference database is widely recognised (Henderson 2007; DeRidder 2011; MacEwan, Angjeli, & Gatenby 2013; Ferguson, Moore, & Schmoller 2015;) and is motivated by the desire to aggregate data about research institutions (such as funding awarded or articles published) to augment analysis and yield insight. Although existing databases seek to provide solutions to this problem (Ringgold¹, ISNI², Orgref³) they either lack global coverage, have restrictive licensing, contain insufficient metadata, or do not meet basic data quality requirements. Hence, Digital Science⁴ has developed GRID to meet the needs of its own internal products, and published the database for free under a CC-BY license to foster innovation, increase interoperability, and promote open standards. GRID was released in October 2015 and is continuing to grow in size and coverage.

POLICIES

One of the first questions to answer when starting on such an endeavour is *what constitutes a research organisation?* GRID’s working definition is an organisation associated with any kind of research activity. This includes awarding grants, receiving grants, submitting patents, publishing articles in journals / conference proceedings, publishing books, publishing journals, publishing datasets or open source software, hosting conferences, participating in clinical trials, etc. Hence, GRID includes a variety of organisation types including universities, hospitals and companies. Only the parent organisation is currently added to GRID, with a small number of exceptions made to cater for national research endeavours such as the Max Planck Society and Chinese Academy of Sciences. It is discipline agnostic and is designed to cover organisations from health, science, engineering, social sciences and the humanities.

A second important issue is where to source candidate organisations from. The approach taken by GRID is pragmatic because it would be impossible to compile and accurately maintain a complete list of all organisations. Instead, it makes practical sense to cover the most frequently occurring instances to maximise the utility of the dataset and minimise curation

¹ <http://www.ringgold.com/>

² <http://www.isni.org/search>

³ <http://www.orgref.org/>

⁴ <https://www.digital-science.com/>

cost. To this end, GRID is built to cover organisations mentioned most often in grant data and article publication sources.

Finally, to achieve consistency within the database, a set of comprehensive policy guidelines is required addressing specific areas that cause confusion or a difference in opinion. GRID uses a team of expert curators who meticulously follow a well defined workflow drawing on corroborative evidence from reputable sources, such as the organisation's own web pages or trusted government web sites.

METADATA AND COVERAGE

GRID currently contains 56,587 organisations with metadata as described in Table 1. The metadata fields were chosen to support:

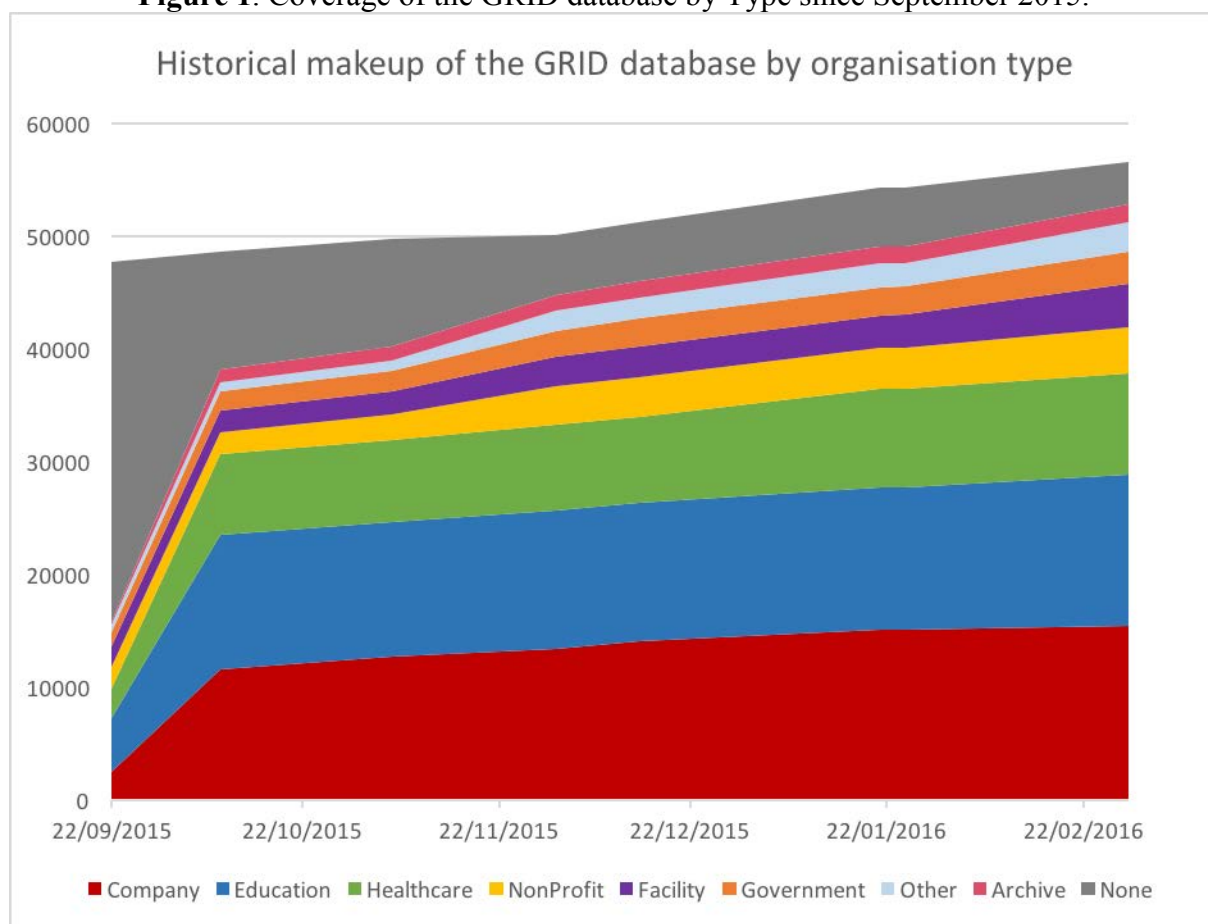
- global geography using an open standard that provides cities, regions and countries as well as precise locations
- a variety of names to support disambiguation
- a high level categorisation of organisation types
- relationships to support parent / child hierarchies and lateral affiliations
- links to as many websites and other id systems as possible

Every record in the GRID database has an address and canonical name (100% coverage), 93% of entries have been assigned a type, and 53% contain links to Wikipedia pages and organisation websites.

GRID has been seeded using mostly North American, European Union and Australasian funding data, as well as public sources of journal article affiliation data and widely used ranking lists. The coverage is global, with entries spanning 214 countries. Up-to-date coverage statistics can found of the GRID website <https://grid.ac/stats>.

Table1. Metadata fields included in GRID and their descriptions.

Field	Description
Name	The canonical name of the organization as defined by itself. In English if possible otherwise in the local language
Type	One of the following options: Company, Education, Healthcare, Nonprofit, Facility, Government, Other, Archive
Address	The city, state and country of the primary address. Defined with and linked to Geonames.
Geo	Expressed as latitude and longitude
Established Date	The year the organisation was established
Closed Date	The year the organisation was dissolved
Links	The organisation's URL(s)
Wikipedia Page	The URL of the corresponding Wikipedia entry
External IDs	Identifiers used in other databases such as ISNI, the CrossRef Open Funder Registry (formally Fundref), Orgref, Wikidata, UKPRN, HESA, UCAS
Aliases	Other alternate names in common use
Labels	Specific language labels (language code & label) in local languages
Relationships	Parents / Children of the organisaiton. Related organisations.

Figure 1. Coverage of the GRID database by Type since September 2015.

The database continues to be populated with new entries and with additional metadata. Figure 1 depicts the increasing coverage of the database broken down by type. The initial sharp increase is due to bulk setting of types using systematic pattern matching. The first public release contained 47,648 entries, with an average increase of 1,787.8 entries per month.

FUTURE WORK

GRID will continue to grow in both size and metadata coverage. Planned extensions to the current data model include the addition of new types to express consortiums and collaborations, mappings to other relevant persistent identification databases, and an overall increase in metadata coverage.

REFERENCES

- DeRidder, J. (2011) Improving the Information Supply Chain with Standard Institutional Identifiers. *Information Standards Quarterly*, 22(3), 26-29.
- Ferguson, N., Moore, R. & Schmoller, S. (2015) Review of selected organisational IDs and development of use cases for the Jisc. *CASRAI-UK Organisational Identifiers Working Group*.
- Henderson, H. (2007) Institutional identifiers and the Journal Supply Chain Efficiency Improvement Pilot. *Serials* 20(3), 180-183.
- MacEwan, A., Angjeli, A., & Gatenby, J. (2013) The International Standard Name Identifier (ISNI): The Evolving Future of Name Authority Control. *Cataloging & Classification Quarterly*, 51(1-3), 55-71. doi: 10.1080/01639374.2012.730601



Submission for the S&T Indicators Conference will be held in València (Spain), 14-16 September 2016, www.sti2016.org

Differential Effects of Scopus vs. Web of Science on University Rankings: A Case Study of German Universities

Wolfram Horstmann, Birgit Schmidt, University of Göttingen, State and University Library

ABSTRACT

Universities usually shape their research agendas based on internal strategic analyses of their strengths and capacities. Yet, external rankings dominate public perception. This constitutes a tension field, especially because external rankings are not predictable. The change from Web of Science to Scopus in the Times Higher Education World University Ranking (THE) between the 14/15 and 15/16 edition presents a prominent example of unpredictability and, simultaneously, provides an opportunity to investigate these effects with respect to the "Citations" component of the indicator set THE applies. In our case study, we analyse these effects for German Universities. We find that effects are strongly differential for individual universities and, even more, not easily reproducible, which limits predictability and transparency of the ranking. We conclude that University rankings should comply with open data principles and universities should take control of their data to start developing their own indicator systems.

In recent years, German universities are increasingly paying attention to the results of international university rankings, e.g. the Times Higher Education (THE) World University Ranking and the Shanghai Academic Ranking of World Universities (ARWU). This type of ranking is typically based on metrics and indicators which reflect the performance of universities in a several areas, for the THE Ranking the five core indicators are teaching, international outlook, research, citations, and industry income [1]. For the 2015-2016 results it has been noted that Germany is the second most-represented nation in the list of top universities in Europe, with 36 institutions, almost a third of which are in the top 50, – after the UK which has topped the ranking of the 200 best universities in Europe, taking nearly a quarter of places (46) [2]. A recent report that assesses the results of the German Excellence Initiative points out that the initiative has made German research more internationally collaborative. Indeed, in the THE Ranking top 100 records an increase of German universities, from 3 in 2010 to 9 in 2015, can be observed (cf. Fig. 1, own plot based on [3]). However, the report also notes that it must be taken into account that during this period the relative weighting of indicators has been changed [4].

Figure 1. From 2011 to 2016 the number of German Universities in the Top 100 has increased from 3 to 9 institutions

It seems that the change of underlying data and methods affected the indicator „citations“ in a rather strong fashion. Overall, 137 institutions (16.7 percent) have seen a decrease in citations from 2015 to 2016, on average by -13.1 points (median: -8). Among these, for only 36 institutions (4.4 percent) the decreases was by more than 15 points. In comparison, 255 institutions (31.2 percent) have seen an increase in citations from 2015 to 2016, on average by 8.3 points (median: 7.3). For 37 institutions (4.5 percent) the increase was by more than 15 points. However, comparing the changes from 2014 to 2015, there were 140 institutions (17.1 percent) with a decrease in citations, on average by -2.5 points (median: -1.5). In comparison, 224 institutions (27.4 percent) have seen an increase in citations from 2014 to 2015, on average by 3.4 points (median: 2.3). Given the strong weighting of citations in the institutional ranking (30%, cf. [1] and Fig. 2) it is therefore not surprising that the 2016 THE Ranking has produced new „winners“ and „losers“.

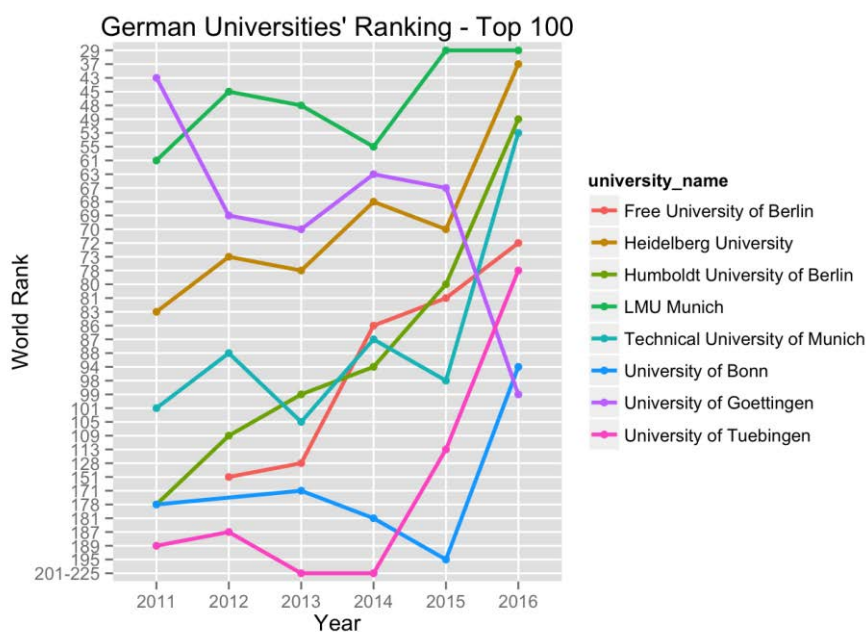


Figure 2: From 2011 to 2016 citation counts changed. Left: Universities with a drop. Right: Changes in Top 15 Universities

Our study has a closer look at effects on mean performance of a sample of 10 German universities resulting from recent changes in the underlying data of the THE Ranking, in particular the move from citations based on Thomson Reuters' Web of Science (until the 2014-2015 rankings) to Elsevier's Scopus database (from the 2015-2016 rankings). A significant effect on citation counts could for example be observed for the University of Göttingen (cf. Fig. 2). Notable changes also include the removal of papers (649 papers) with more than 1,000 authors from the citations indicator (a step which has been questioned by the scientific community, cf. [5]). We further investigate effects on German universities' performance in the THE Ranking and assess possible explanations. In particular, effects on citation counts and the representation of publications of these universities in both databases, the impact of measuring citations of English language publications, and the removal of publications with more than 1,000 authors (which primarily result from High Energy Physics research) will be taken into account.

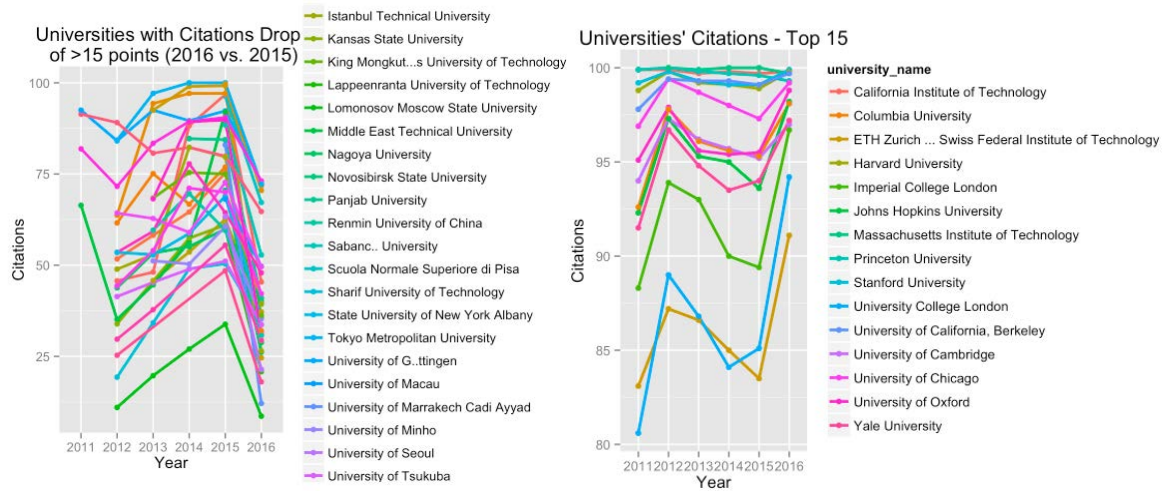
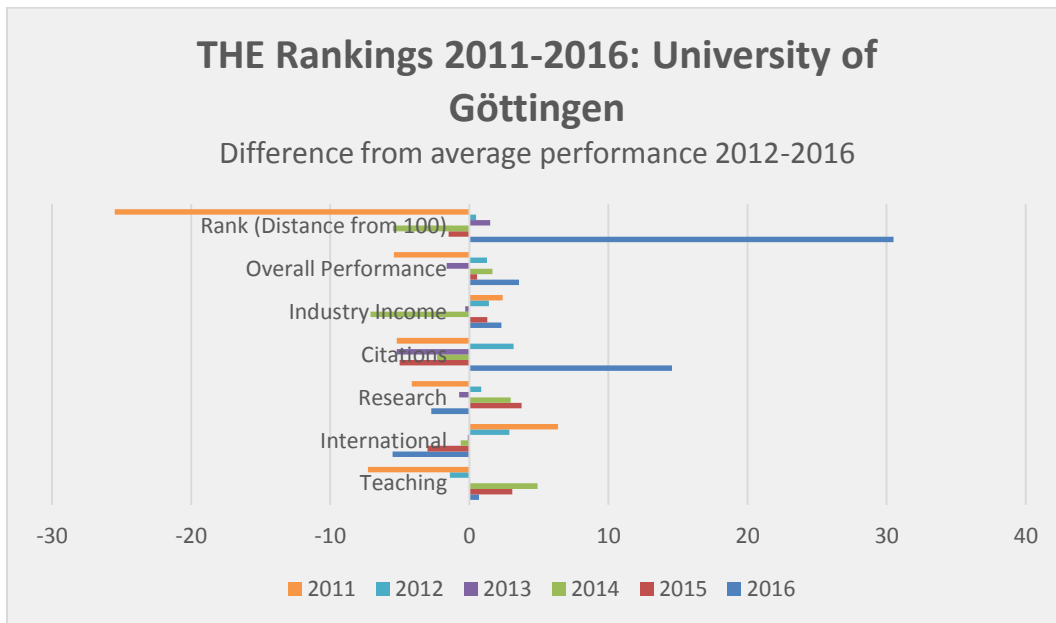


Figure 2. Mean difference of Göttingen University’s performance in the THE Ranking indicators.



REFERENCES

- [1] World University Rankings 2015-2016 methodology, 24 September 2015, <https://www.timeshighereducation.com/news/ranking-methodology-2016>
- [2] U. Brandenburg: Best universities in Europe 2016: secrets of Germany's success, 10 March 2016, <https://www.timeshighereducation.com/world-university-rankings/best-universities-in-europe-2016>
- [3] Times Higher Education (THE) World Universities Ranking data retrieved from Kaggle on 20 March 2016, <https://www.kaggle.com/mylesoneill/world-university-rankings>
- [4] Internationale Expertenkommission (IEKE): Evaluation der Exzellenzinitiative, Endbericht, Januar 2016, <http://www.gwk-bonn.de/fileadmin/Papers/Imboden-Bericht-2016.pdf>
- [5] <https://www.timeshighereducation.com/blog/world-university-rankings-blog-dealing-freak-research-papers>



On the normalization of citation impact based on the Essential Science Indicators classification of Thomson Reuters

Olga Baranova* and Alfred Peris*

*olba@upvnet.upv.es; aperis@mat.upv.es

Instituto de Matemática Pura y Aplicada, Universitat Politècnica de València, Camí Vera S/N, València, 46022 (Spain)

ABSTRACT

In recent years the studies on the normalization of citation counts (or citation impact) has attracted the interest of many researchers. Several algorithms and procedures have been proposed in order to normalize citation impact, usually based on classification systems or on fractional citation weighting. It is clear that publication rates and citation behaviour vary considerably from field to field, they depend on period of time, etc., which makes the necessity of understanding the different behaviours, and to establish normalization procedures. The Essential Science Indicators (ESI) classification provides 22 subject areas in science and social sciences and it is based on journal assignments. Each journal is assigned to only one of the 22 subject areas, and the overlap between subject areas is 0. This fact makes the normalization based on the ESI classification especially simple and easy to use.

Concerning indicators, we perform normalization to both types, size-dependent and size-independent ones. That is, to indicators that depend on the number of publications (size-dependent), and to indicators that do not depend on the number of publications (size-independent). It is also important to normalize with respect to the year of publication. We do so by considering average citation rates and percentiles. The results are illustrated with certain sets of publications, corresponding to individuals and to a research unit.

The citation rates, as yearly averages of citations per paper, provide us size-independent indicators. The baseline percentiles define levels of citation activity in such a way that the larger the minimum number of citations for the corresponding year, field and percentile, the smaller the peer group, thus representing size-dependent indicators.

When comparing individuals, we can perform both: The normalization using citation rates (size-independent) and baseline percentiles (size-dependent), and taking into account year of publication and subject area. When comparing different research units, we can use the two types of normalization procedures too. Finally, a comparison of individuals with research units (or larger groups of people) can only be done through size-independent indicators, that is, by using the citation rates in our case of study.

REFERENCES

- Colliander, C. & Ahlgren, P. (2011). The effects and their stability of field normalization baseline on relative performance with respect to citation impact: A case study of 20 natural science departments. *Journal of Informetrics*, 5, 101-113.
- Perianes-Rodriguez, A. & Ruiz-Castillo, J. (2016). A comparison of two ways of evaluating research units working in different scientific fields. *Scientometrics*, 106, 539-561.
- Waltman, L. & van Eck, N.J. (2013). A systematic empirical comparison of different approaches for normalizing citation impact indicators. *Publications*, 1, 78-86.
- Watman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10, 365-391.
- (2014). InCites Indicators Handbook from:
<http://ipscience-help.thomsonreuters.com/inCites2Live/indicatorsGroup/aboutHandbook/appendix/indicatorsTables.html>



Rock around the clock? Exploring scholars' downloading patterns

Sarah Cameron-Pesant*, Yorrick Jansen** and Vincent Larivière***

* sarah.cameron-pesant@umontreal.ca

Université de Montréal, École de bibliothéconomie et des sciences de l'information, 3150 Jean-Brillant,
H3T 1N8 Montréal, Qc. (Canada)

** yorrick.jansen@gmail.com

1science, 3863 St-Laurent Blvd. Suite 206, H2W 1Y1 Montréal, Qc. (Canada)

*** vincent.lariviere@umontreal.ca

Université de Montréal, École de bibliothéconomie et des sciences de l'information, C.P. 6128, Succ. Centre-Ville, H3C 3J7 Montréal, Qc. (Canada) and Université du Québec à Montréal, Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Observatoire des Sciences et des Technologies (OST), C.P. 8888, Succ. Centre-Ville, H3C 3P8 Montréal, Qc. (Canada)

ABSTRACT

In the context of the current “publish or perish” culture of academe, scholars' working habits and work-life balance issues have become a subject of interest. The aim of this paper is to examine scholars' downloading patterns on four different timescales: daily, weekly, monthly and by season. Our analyses are based on the web log data of *Érudit*, the main journal diffusion platform of French-Canadian journals in the social sciences and humanities. We used a stratified random sample of 70 days of log files (total of 1,407,163 downloads) from 2011 to 2015. We focus on the countries with the highest number of downloads: Canada, France and the United States. Results show that Canadian and French users perform the large majority of their downloads during weekdays. Americans' online activity, however, is steadier across the day (and night) and during weekends. Canadian users' download activity increases at the middle of semesters, probably due to the high number of undergraduates using *Érudit*.

CONTEXT

The issue of work-life balance has been discussed since the industrial revolution (Guest, 2002, p. 256). However, as stated by Guest (2002), it “has come to the fore in contemporary debates largely because in affluent society the excessive demands of work are perceived to present a distinctive issue that needs to be addressed” (p. 256). The pressures of work are getting heavier in modern societies, and scholars are no exception, especially in the current “publish or perish” culture. In this context, many authors have investigated researchers' working habits, such as work-life balance in information science (Cabanac & Hartley, 2013), calendar effects on the dissemination of science (Magnone, 2013), researchers' timetable (Wang & al., 2012), and seasonal influences and academic life cycles from a year to another (Moed & Halevi, 2016). This brief communication aims at contributing to this literature by analysing the log data of *Érudit*, the main journal diffusion platform for French-Canadian journals. More specifically, it examines users' downloading behaviour on four timescales: daily, weekly, monthly and by season.

METHODS

Our study relies on *Érudit*'s web log data. This journal diffusion platform provides access to (mostly) French-Canadian journals in the social sciences and humanities, comprising 119 scholarly journals. We used a stratified random sample of 70 days of *Érudit*'s log files for five years, from 2011 to 2015. Overall, 1,407,163 downloads are analysed in this paper. In order to analyse the time patterns of downloads by country, the time zone of users was taken into account. A Python script was used to parse the data and identify successful downloads of scholarly papers. A robot detection technique was implemented to exclude downloads performed by web crawlers and robots behaving like humans. The geolocation was found with the IP address of the users, and the parsed web logs were imported in a PostgreSQL database.

RESULTS

Daily and weekly working patterns

A first set of analyses examined the proportion of downloads made by time of day (Figure 1) as well as by day of week (Figure 2). Figure 1 reveals that Canadian researchers start their day slightly earlier than the French, but that the French continue to work later in the afternoon and early evening. For both countries, we see the effect of lunch time (a longer and deeper decline in France) and dinner (two hours later in France). However, as was shown by Wang and his colleagues (2012), American researchers follow a very different timetable: they are active at night, even during the weekends. Their activity is steadier across the day (and night), suggesting that they eat at different moments or that they continue working while eating.

Figure 1: Proportion of downloads made by time of day, for Canada, France and USA.

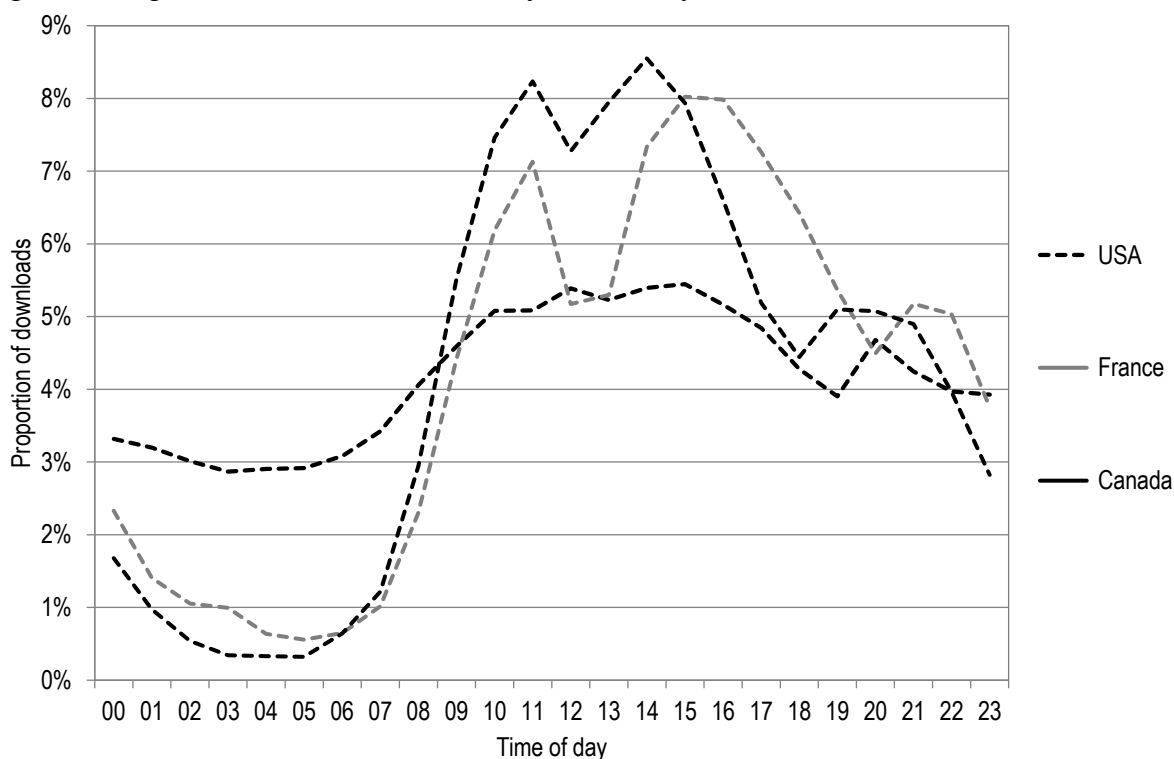
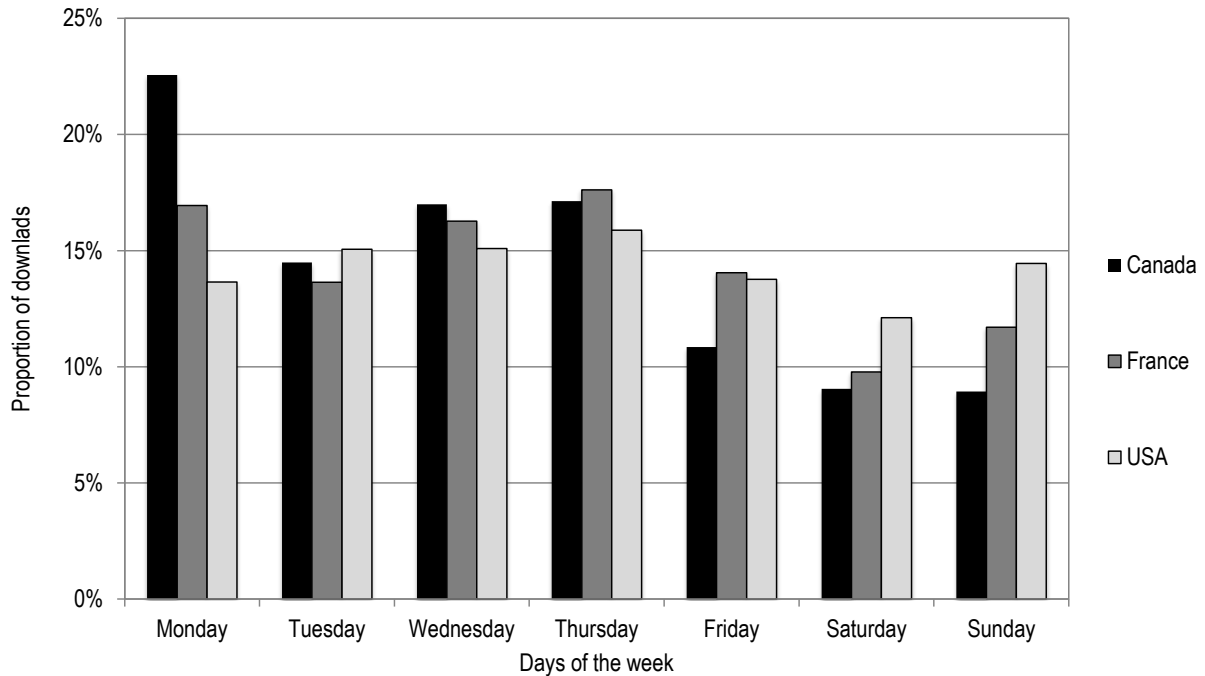


Figure 2 presents the weekly download patterns. Our results are consistent with those of Wang and his colleagues (2012): the overall number of downloads decreases in the weekends, at least for Canada and France. Magnone (2013) and Cabanac and Hartley (2013) also observed a decline in the number of paper submissions during weekends. Monday seems to be

the busiest day for Canada and France. Once again, the American average is steady across the week. We even observe for US scholars a higher proportion of downloads made on Sunday than on Friday.

Figure 2: Proportion of downloads by day of the week, for Canada, France and USA.



Downloads throughout the academic year

Figure 3 shows the monthly variations in downloads. Canadian users are especially active from September to November and from February to March—therefore mainly in the middle of semesters, which suggests intensive use by undergraduate students. In Canada and France, a decrease in downloads for December and January can likely be explained by holidays. The decrease that starts in March might be influenced by special events like conferences or spring break. Figure 4 reveals that the daily average of downloads for the three countries goes down steadily as the academic year goes by (fall being the highest and summer the lowest).

Figure 3: Proportion of downloads by month, for Canada, France and USA.

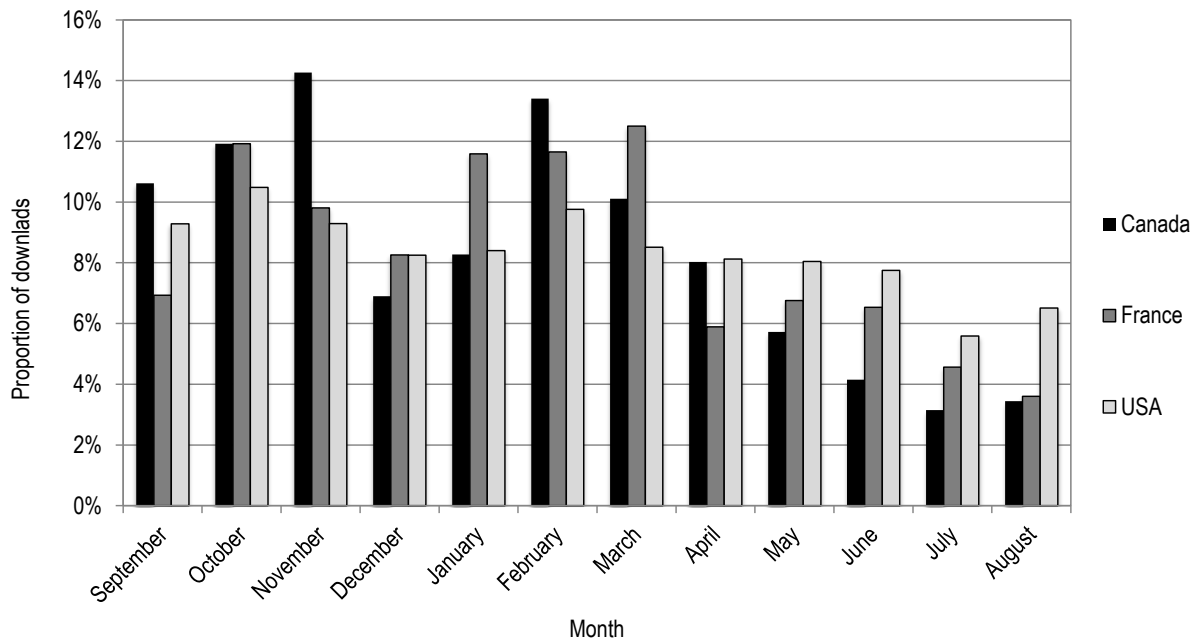
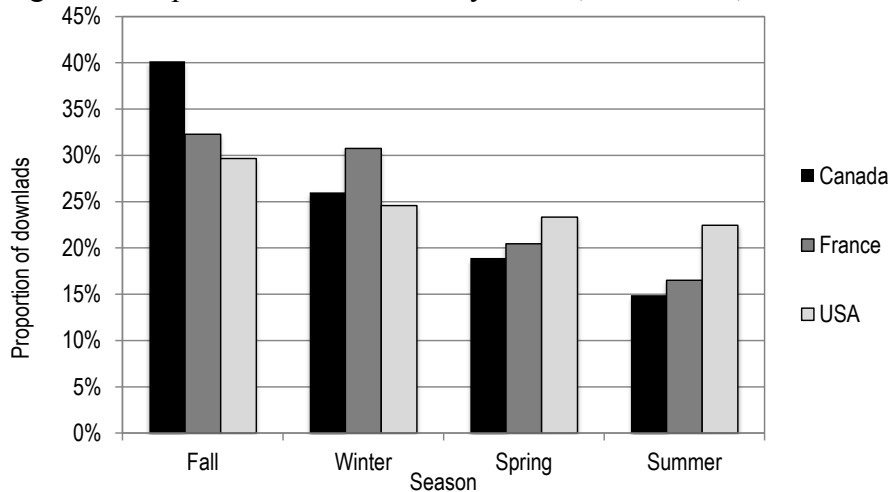


Figure 4: Proportion of downloads by season, for Canada, France and USA.



DISCUSSION

On the whole, our results on researchers' downloading patterns are consistent with those obtained by other studies. Cabanac and Hartley (2013) saw, for instance, an increase in the researcher's work-related online activity between 2001 and 2012. However, Magnone (2013) found that the number of submissions was lower in fall and winter, which does not converge with our results. This is likely due to the fact that, while paper submissions provide an indicator of researchers-authors working habits, downloads can be performed by a much broader set of users, such as undergraduates, practitioners or the general public (Moed & Halevi, 2016). Along these lines, the lower number of submissions observed by Magnone (2013) might be due to the fact that fall and winter are more teaching-intensive than the two other seasons.

One of the most surprising results we found is the difference between the activity of American users and that of Canadian and French users. The steadiness of Americans'

downloading habits, regardless of the time scale we look at, is stunning. This might be due, again, to different types of *Érudit* users found in the various countries. While most of *Érudit* users in Canada and Quebec might be students—as this is one of the most important source of French scholarly information—downloads made in France and, especially, in the United States, might be performed by researchers, which would explain the higher proportion of downloads made in the summer.

REFERENCES

- Cabanac, G., & Hartley, J. (2013). Issues of work–life balance among JASIST authors and editors. *Journal of the American Society for Information Science and Technology*, *64*(10), 2182-2186. doi:10.1002/asi.22888
- Guest, D. E. (2002). Perspectives on the study of work-life balance. *Social Science Information*, *41*(2), 255-279. doi:10.1177/0539018402041002005
- Magnone, E. (2013). A scientometric look at calendar events. *Journal of Informetrics*, *7*, 101-108. doi:10.1016/j.joi.2012.09.006
- Moed, H. F., & Halevi, G. (2016). On full text download and citation distributions in scientific-scholarly journals. *Journal of the Association for Information Science and Technology*, *67*(2), 412-431. doi:10.1002/asi
- Wang, X., Xu, S., Peng, L., Wang, Z., Wang, C., Zhang, C., & Wang, X. (2012). Exploring scientists' working timetable: do scientists often work overtime? *Journal of Informetrics*, *6*(4), 655-660. doi:10.1016/j.joi.2012.07.003



Research leadership in key fields for emerging and developing countries

Gregorio González-Alcaide^{*}, Charles Huamani, Jinseo Park^{**} and José M. Ramos^{**}

^{*}gregorio.gonzalez@uv.es

Department of History of Science and Documentation, University of Valencia, 15 Blasco Ibáñez Avenue, Valencia, 46010 (Spain).

^{**}huamani_ca@hotmail.com; jayoujin@hanmail.net; jramosrincon@yahoo.es

Instituto Nacional de Salud, 1154 Av. Arriba Peru, 42 (Perú); Korea Institute of Science and Technology Information (KISTI), 66 Hoegiro, Dongdaemun-gu, Seoul, 130-741 (South Korea); Department of Medicine, Universidad Miguel Hernández de Elche, Avenida de la Universidad s/n, Elche, 03202 (Spain) and Department of Internal Medicine, Hospital General Universitario de Alicante, 11 Pintor Baeza, Alicante, 03010 (Spain).

ABSTRACT

Rules for the order of multiple authors in a document are generally consistent within a field. In medicine, the first author is the main contributor in terms of involvement and/or leadership in the work, while the contributions of subsequent authors have successively less weight. The corresponding authors and final authors are also considered to have a prominent role among all signatories of the publication. The objective of the present study is to analyze the order of the signatures and the address for correspondence as indicators for assessing North-South relationships and leadership in scientific research. A bibliographic search was carried out in Science Citation Index-Expanded database, identifying all articles and reviews published during the 2011-2015 period in the categories of Infectious Diseases, Parasitology, Tropical Medicine and Paediatrics. Countries were categorized according to their Human Development Index (HDI). The analysis shows an asymmetrical relationship between countries in the North and South with regard to leadership and citation rates in scientific publications. In collaborative papers, the analysis of the order of signatures may constitute an indicator to measure the leadership of research.

INTRODUCTION

Authorships are the mechanism through which scientists assume responsibility for published content and take credit for new ideas or discoveries. Rules for the order of multiple authors in a document are generally consistent within a field. In medicine, the first author is the main contributor in terms of involvement and/or leadership in the work, while the contributions of subsequent authors have successively less weight. The corresponding authors and final authors are also considered to have a prominent role among all signatories of the publication. In collaborative papers, the analysis of the order of signatures may be used as an indicator to measure leadership in research (Kim, 2006; Schubert & Sooryamoorthy, 2010).

The objective of the present study is to analyse the order of the signatures and the address for correspondence as indicators for assessing North-South relationships according to their Human Development Index (HDI) and leadership in scientific research.

METHODS

A bibliographic search was carried out in Science Citation Index-Expanded database, identifying all articles and reviews published in 2011-2015 in the categories of Infectious

Diseases, Parasitology, Tropical Medicine and Paediatrics. Infectious and parasitic diseases disproportionately affect the least developed countries; for example, they represent 60% of deaths in Africa compared to only 5% in Europe. Tropical Medicine is also of special interest because many developing countries are between the two tropics, where the climatic conditions exist for the development of specific diseases, often considered neglected diseases. Finally, regarding Paediatrics, less developed countries have the highest rates of both stillbirth and infant and child mortality, so basic and clinical research on diseases affecting the pre-adult population should also be a priority in these countries.

Countries were categorised according to their Human Development Index (HDI) into very high human development (VH), high human development (H), medium human development (M) and lower human development (L). The HDI is a measure published by the United Nations Development Programme of average achievement in key dimensions of human development. For each of these four groups of countries, the order of signatures was analysed. Furthermore, the documents were classified in the categories detailed in table 1.

Table 1. Collaboration types considering the order of signatures in scientific publications group by countries' HDI.

First position	Second and subsequent positions	Description
VH or H	-	A single country of very high or high economic development.
VH or H	VH or H	Collaboration between two or more countries of very high or high economic development.
M or L	-	A single country of medium or low economic development.
M or L	M or L	Collaboration between two or more countries of medium or low economic development.
VH or H	M or L	Leadership of a country of very high or high economic development and participation of one or more countries of medium or low economic development.
M or L	VH or H	Leadership of a country of medium or low economic development and participation of one or more countries of very high or high economic development.
VH or H	VH or H + M or L	Leadership of a country of very high or high economic development with simultaneous participation of other countries of very high or high economic development and medium or low economic development.
M or L	M or L + VH or H	Leadership of a country of medium or low economic development with simultaneous participation of other countries of medium or low economic growth and high or very high economic development.

RESULTS

The analysis of the role of leadership by countries' development index (table 2) reveals the dominance of the VH and H countries and the limited importance of other countries, which is especially striking in Paediatrics. H countries contribute fewer signatures compared to VH countries but exercise a prominent leadership role, with much higher percentages in the first position and as corresponding authors in relation to the number of total signatures. Also noteworthy is the limited leadership role of M and L countries, well below the already reduced scientific contribution if measured in number of total signatures.

Table 2. Distribution of total number of signatures, first position and corresponding author in scientific publications group by countries' HDI

Research Area		VH (%)	H (%)	M (%)	L (%)
Tropical Medicine	Total signatures	46.27	26.85	12.93	13.93
	1 st position	38.45	38.92	13.33	9.28
	Corresponding author	40.23	38.65	12.77	8.35
Parasitology	Total signatures	61.25	22.61	8.4	7.74
	1 st position	58.21	29.27	7.54	4.97
	Corresponding author	59.04	29.22	7.12	4.61
Infectious Diseases	Total signatures	71.61	14.3	7.38	6.7
	1 st position	74.44	16.38	5.65	3.52
	Corresponding author	75.45	16.06	5.34	3.15
Pediatrics	Total signatures	81.25	12.19	5.25	1.31
	1 st position	80.83	13.1	5.26	0.8
	Corresponding author	81.17	12.91	5.17	0.74

The analysis of the degree of citation regarding collaboration types (table 3) reveals that the most cited work is that led by VH or H countries with the participation of M or L countries, overtaking the degree of citation of publications led by M or L countries and involving VH or H countries. Publications in which only M or L countries are involved have a lower degree of citation than papers that involve only H or VH countries.

Table 3. Average citations per paper in scientific publication group, by collaboration types

First position	Second and subsequent positions	Tropical Medicine	Parasitology	Infectious Diseases	Pediatrics
VH or H	-	3.49±6.59	6.81±10.76	6.69±23	3.37±6.53
VH or H	VH or H	4.22±6.97	6.7±11.57	6.98±13.59	4.16±7.93
M or L	-	2.4±4.57	3.77±6.24	3.18±5.17	1.83±3.35
M or L	M or L	2.65±4.51	3.78±5.34	3.46±5.43	1.89±3.48
VH or H	M or L	5.52±9.93	5.81±9.89	7.31±11.99	3.71±5.06
M or L	VH or H	4.79±8.7	5.72±9.66	6.25±10.14	3.04±4.45
VH or H	VH or H + M or L	6.56±10.28	7.16±11.3	8.37±14	4.76±7.7
M or L	M or L + VH or H	5.28±7.48	5.82±10.64	6.35±10.14	3.51±5.18

CONCLUSION

The analysis shows an asymmetrical relationship between countries in the North and South with regard to leadership (first position in the order or signatures and corresponding authors) and citation rates in scientific publications.

REFERENCES

Kim, K. W. (2006). Measuring international research collaboration of peripheral countries: taking the context into consideration. *Scientometrics*, 66, 231-240.

Schubert, T., & Sooryamoorthy, R. (2010). Can the centre–periphery model explain patterns of international scientific collaboration among threshold and industrialised countries? The case of South Africa and Germany. *Scientometrics*, 83, 181-203.



Mass Gathering as an emerging field: a co-citation analysis

Gregorio González-Alcaide^{*}, Pedro Llorente^{**} and José M. Ramos^{**}

^{*}gregorio.gonzalez@uv.es

Department of History of Science and Documentation, University of Valencia, 15 Blasco Ibáñez Avenue, Valencia, 46010 (Spain).

^{**}pedrollorente@hotmail.com; jramosrincon@yahoo.es

Center for Development Cooperation and Volunteerism, Universidad Miguel Hernández de Elche, Avenida de la Universidad s/n, Elche, 03202 (Spain); Department of Medicine, Universidad Miguel Hernández de Elche, Avenida de la Universidad s/n, Elche, 03202 (Spain) and Department of Internal Medicine, Hospital General Universitario de Alicante, 11 Pintor Baeza, Alicante, 03010 (Spain).

ABSTRACT

A mass gathering is as a planned or spontaneous event where the number of people attending is sufficient to strain the planning and response resources of the community or nation hosting the event. The aim of this study is to analyze the development of the emerging field of mass gathering through the study of the scientific publications on the topic. A bibliographic search was carried out in Science Citation Index-Expanded, identifying all papers published using the term “Mass gathering”. We analyze the intellectual structure of the area through a co-citation analysis, identifying the countries with the largest presence in the group of most frequently cited core documents. We also analyzed the age of the cited bibliography in publications on mass gatherings in order to determine the usefulness of this indicator in identifying the existence of an emerging topic. The mass gatherings literature has experienced a significant development since 2006. The performance of the co-citation analysis revealed five clusters or groups of documents polarized around two large thematic clusters: infectious diseases associated with pilgrimages to Mecca and health services planning and response in relation to large sporting events. The comparative analysis of the comprehensive contribution per country to research on mass gatherings and the contribution of those same countries to the group of most frequently co-cited documents reveals that the most productive countries increase the relative weight of their contributions to the group of core documents. In contrast, the least developed countries are almost entirely unrepresented within the group of core documents. The fact that the bibliography cited in the publications on mass gatherings is more recent compared to the age of all publications in the primary areas of knowledge related to the topic, underlines that this indicator may be useful in identifying and monitoring the development of an emerging topic.

INTRODUCTION

The World Health Organization (WHO) defines a mass gathering as a planned or spontaneous event where the number of people attending is sufficient to strain the planning and response resources of the community, state or nation hosting the event (World Health Organization, 2008). Mass gatherings provide ideal conditions for the transmission of different diseases, and they can generate a risk of situations like human avalanches, so planning and surveillance measures are essential to ensure participants' and populations' health wherever mass gatherings occur (Black et al., 2014).

Bibliometrics allows, by means of methods such as co-citation analysis, to delineate the main lines of research under study in a particular discipline and to determine the documents that have played the largest role in laying the foundation for its development.

The aim of this study is to analyze the development of the emerging field of mass gathering through the study of the scientific publications on the topic, identifying the works of reference that constitute the intellectual basis for current knowledge on mass gathering as well as the countries responsible for producing them.

METHODS

A bibliographic search was carried out in Science Citation Index-Expanded database, identifying all papers published until 2015 using the term “Mass gathering” (topic field). We analyze the size and stability of the research community involved in the research on the topic. For this, we calculated the percentage of transient authors (those who have published only one paper) and the big producers (authors with more than nine published papers). In order to analyze the intellectual structure of the area, we perform a co-citation analysis, identifying the contribution per country, comprehensively for the group of documents as a whole, and specifically with regard to the core co-cited documents. Moreover, we analyze the age of the cited bibliography in publications on mass gatherings in order to determine the use of this indicator in establishing the existence of an emerging topic.

RESULTS

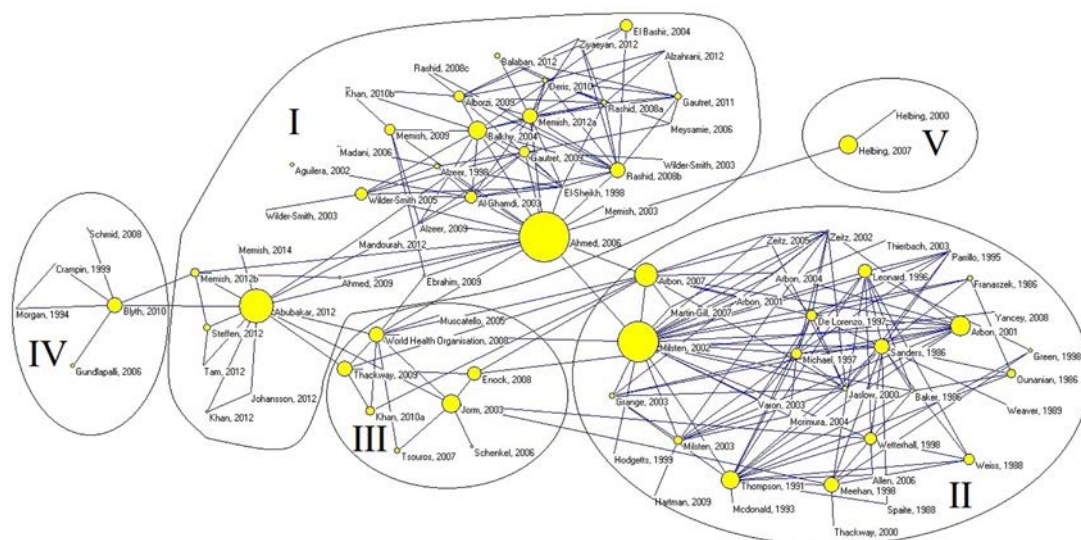
A total of 278 documents on mass gatherings were identified from the search in the Web of Science database: 185 articles (66.55%), 28 reviews (10.07%), 24 commentaries or editorials (8.63%), 22 proceedings papers (7.91%), 11 letters (3.96%) and 8 documents of other types (2.88%). The evolution in the number of documents published per year shows a dramatic increase in the number of works published in the 2006–2010 period, when 263% more documents were published than in the previous five-year period. This growth continues into the next five-year period (2011–2015), when the growth rate stands at 135%

The documents identified were published by 1050 different authors, of whom 85.62% (n=899) were transient authors, that is, they published a single document. Another 14.09% (n=148) signed 2 to 9 papers, while only three authors (0.18%) published 10 or more: Z. A. Memish (Ministry of Health and College of Medicine, Alfaisal University, Riyadh, Saudi Arabia), M. Barbeschi (Department of Communicable Disease Surveillance and Response, WHO, Geneva, Switzerland) and J.A. Al-Tawfiq (Johns Hopkins Aramco Healthcare, 203 Dhahran, Saudi Arabia).

Of the 278 documents identified through the Web of Science, 251 included a total of 7149 bibliographic references (mean 28.48 ± 24.63). These references were used to generate a co-citation matrix of documents, made up of 165,899 different pairs of bibliographic references. The overwhelming majority of the co-citations (96.56%, n=160,196) occurred only once, 2.63% (n=4366) appeared twice, 0.47% (n=787) appeared three times, 0.18% (n=300) appeared four times, and 0.15% (n=250) of the co-citations appeared in five or more documents. 87 documents had the highest degree of co-citation, which constitute the most prominent sources of reference in terms of the degree of citation and the linkage with other documents (core-documents).

The performance of a clustering algorithm revealed five clusters or groups of documents (figure 1). Most of the papers appear to be polarized around two large thematic clusters. One of them (I), bringing together 36 papers, is made up of studies that analyze the prevalence of infectious diseases associated with pilgrimages to Mecca, particularly influenza but also other pathologies like tuberculosis, pneumonia, pertussis, meningitis, and meningococcal disease. The other main thematic cluster (II), comprising 35 documents, deals with health services planning and response in relation to large sporting events like the Olympic Games and the FIFA World Cup. Another thematic cluster (III, n=8) is made up of studies that highlight the importance of developing public health surveillance systems, particularly in countries or cities that are planning to host events attracting large numbers of visitors. Finally, two other small clusters are dedicated to analyzing the incidence of infections caused by *Escherichia coli*, *Campylobacter*, or influenza at events like music festivals (IV, n=5); or developing simulation models associated with crowd behavior during or immediately after catastrophes, such as panic induced human avalanches, with the object of better understanding, responding to, and mitigating the consequences of these situations (V, n=2).

Figure 1. Main research clusters identified through the co-citation analysis of publications on mass gatherings included in the Web of Science.



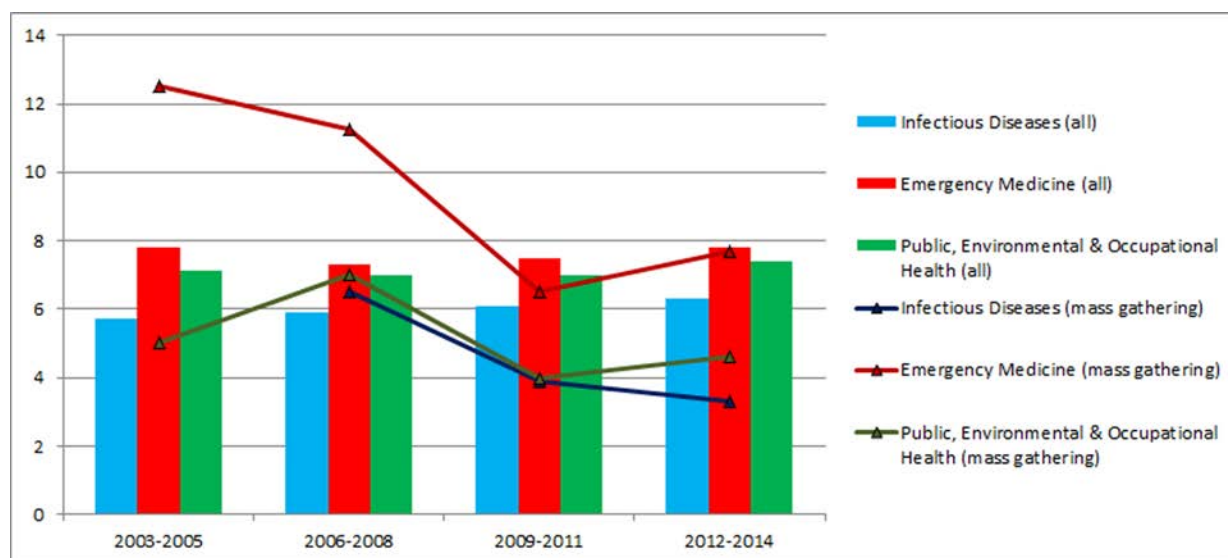
The participation in science and in scientific publications should be measured at different levels in order to evaluate both the comprehensive contributions to the body of scientific publications on a topic, and the most significant contributions toward the advancement of knowledge. The differences between these two facets may be sizeable. For example, the comparative analysis of the comprehensive contribution per country to research on mass gatherings and the contribution of those same countries to the group of most frequently co-cited documents (core documents) reveals that the most productive countries increase the relative weight of their contributions to the group of core documents (table 1), that is, their position becomes even more central with regard to their contribution to knowledge in the field. In contrast, the least developed countries, which contribute fewer documents to the body of publications as a whole, are almost entirely unrepresented within the group of core documents.

Table 1. Distribution of scientific contributions to the area of Mass Gathering, by country

Country	% of contributions to the field	
	All documents	Core co-cited documents
USA	34.51	40.68
UK	18.82	23.73
Saudi Arabia	18.43	42.37
Australia	13.33	11.86
France	7.84	10.17
Switzerland	7.84	13.56
Germany	6.67	3.39
Canada	4.71	11.86
Sweden	4.31	0
Italy	2.74	0
Greece	2.35	0
Iran	2.35	5.08
India	1.96	0
Japan	1.96	0
People's R China	1.57	0
South Africa	1.57	3.39
Brazil	1.18	0
Denmark	1.18	0
Hungary	1.18	1.69
Jordan	1.18	0
Malaysia	1.18	1.69
Mexico	1.18	1.69
Poland	1.18	0
Singapore	1.18	5.08
Spain	1.18	0
Taiwan	1.18	0

The analysis of the age of the bibliographic references cited in documents about mass gatherings falling under the three main fields feeding into the generation of knowledge on the topic (Infectious Diseases; Emergency Medicine; and Public, Environmental & Occupational Health) (figure 2) reveals that the cited bibliography is more recent in publications on mass gatherings. Our study reveals an additional nuance related to this indicator, suggesting that this phenomenon does not take place—or takes place at a much smaller scale—during the initial stages of a topic's development. Indeed, the earliest papers on a given topic may draw from an older body of work, and only when momentum in the field begins to pick up will the average age of references drop.

Figure 2. Evolution of the mean age of the cited bibliography of publications on mass gatherings included in the Web of Science database.



CONCLUSIONS

The mass gatherings literature has experienced a significant development since 2006. Most of the papers appear to be polarized around two large thematic clusters: infectious diseases associated with pilgrimages to Mecca and health services planning and response in relation to large sporting events.

The central-peripheral relationship between countries with higher economic and research development and less developed countries demonstrates a more pronounced asymmetry when considering the most influential documents that establish the intellectual basis and knowledge of reference in the field. The integration of peripheral countries should be promoted, not only in scientific activities but also among core documents.

One last aspect of interest to keep in mind in the present study is that the age of the cited bibliography in the documents may be an indicator of reference for determining the existence of an emerging topic of investigation. Given the more recent nature of the bibliography cited in the publications on mass gatherings, this aspect should be compared between mass gatherings and other emerging topics (Jarić et al. 2014).

REFERENCES

Black, G., Kononovas, K., Taylor, J., & Raine, R. (2014). Healthcare 497 Planning for the Olympics in London: A Qualitative Evaluation. *PLoS ONE*, 9, e92338.

Jarić, I., Knezević-Jarić, J., & Lenhardt, M. (2014). Relative age of references as a tool to identify emerging research fields with an application to the field of ecology and environmental sciences. *Scientometrics*, 100, 519-529.

World Health Organization (WHO). (2008). *Communicable disease alert and response for mass gatherings*. Geneva: World Health Organization. Retrieved March 26, 2016 from: http://www.who.int/csr/resources/publications/WHO_HSE_EPR_2008_8c.pdf?ua496=1



Research Activity Classification based on Time Series Bibliometrics

Takahiro Kawamura*, Yasuhiro Yamashita* and Katsuji Matsumura*

*takahiro.kawamura@jst.go.jp; yasuhiro.yamashita@jst.go.jp; matsumur@jst.go.jp
Japan Science and Technology Agency, Tokyo (Japan)

ABSTRACT

Bibliometrics, such as the number of papers cited and frequency, are often used to compare researchers based on specific criteria. The criteria, however, are different in each research domain and are set by empirical laws. Moreover, there are arguments, such as that the simple sum of the metric values works to the advantage of the elders. Therefore, this paper attempts to constitute features from the time series data of bibliometrics and then classify the researchers accordingly. In detail, time series patterns are extracted from a large amount of bibliographic datasets, and then a model to classify whether the researchers are "distinguished" or not is created by using machine learning techniques. The experiments achieved F-measures of more than 80% in the classification of 114 researchers in two research domains based on the datasets of the Japan Science and Technology Agency and Elsevier's Scopus. In the future, we will conduct verification on a number of researchers in several domains, and then we will make use of discovering "distinguished" researchers who are not yet widely known.

INTRODUCTION

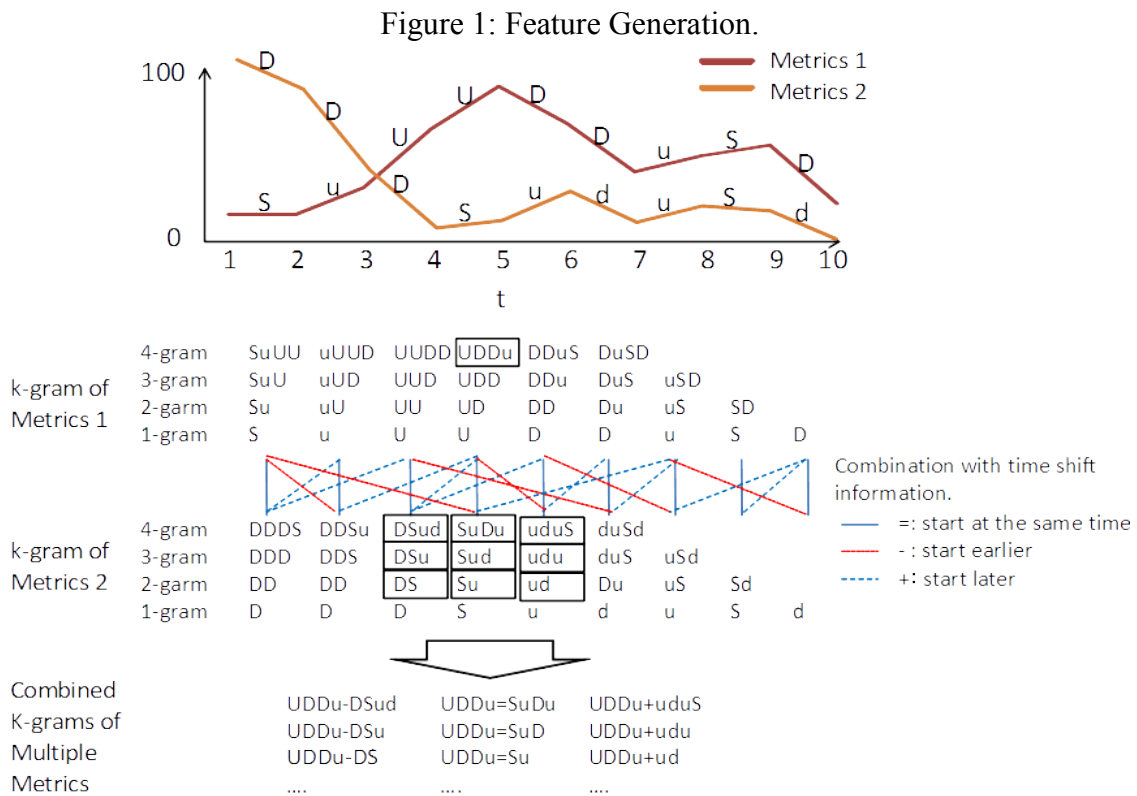
In this paper, we aim to discover "distinguished" researchers who have not been widely known based on bibliometrics. In many cases, some criteria concerning the amount of research achievements, such as the number of papers published and citations received, are determined, and then the research activities are evaluated regarding whether the criteria are satisfied. However, the simple sum of the achievements is consequently beneficial to the elder researchers, and there is a problem in that researchers who have a significant achievement cannot be differentiated from those publishing a few papers over a long period of time. Also, it has been found that the transitions of the achievements of "distinguished" researchers exhibit some patterns (Bjork, Offer & Soederberg 2014). Therefore, this paper finds characteristic patterns from the time series changes of the international and domestic research achievements of "distinguished" researchers, and then it attempts to classify the researchers.

CLASSIFICATION USING TIME SERIES METRICS

Feature Generation

There are several ways to represent time series data as features, such as numerical values and item pairs of an attribute and its value. In this paper, we convert the sequential data of real values to characters in order to reduce the data size. Symbolic Aggregate Approximation (SAX) (Patel et al, 2002) is a well-known method for this purpose. We used the SAX but converted the differences between the two values to represents the changes in achievements instead of the values of the sequential data in the SAX. We also used k-gram (consecutive k characters) in Natural Language Processing (NLP), and then extended it to the combined k-grams that have anteroposterior relations in time, to represent loosely the multiple overlapping sequences.

Figure 1 shows the workflow of feature generation. First, the numerical values are normalized to [0–100] per person, and then differences between consecutive years are converted to six symbols [U, u, S, d, D, 0] to represent changes in time as features, where U = over a 30 point increase from the last year, u = 5–0 point increase, S = +/-5 point change including no change, d = between -5 and -30 point decrease, D = over a -30 point decrease, and 0 = no paper/cited. Then, we generate a k-gram (k = 1–4) of the symbols, where k = 4 indicates a five-year period, and, finally, the k-grams for the five metrics in the following section are combined with anteroposterior relations (+, =, -), comparing the start time t of two time-series, where + indicates that the start time of the following k-gram is after the start time of the first k-gram, = indicates that the start time is the same as the first k-gram, and - indicates that the start time is before the first k-gram.



Feature Selection

Next, to find characteristic patterns from a set of patterns created in the previous section, which are important features in the following machine learning phase, we calculate P_{ij} (Equation 1) for each pattern, which corresponds to Term Frequency/Inverse Document Frequency (TF/IDF) in NLP. P_{ij} filters the general patterns that are common with several researchers. Finally, we selected the patterns of the ten highest P_{ij} per person. Thus, if there are 50 persons, at most 500 patterns become features after deleting duplicated patterns.

$$P_{ij} = paf_{ij} \times \log\left(\frac{N}{pef_j}\right) \quad (1)$$

paf_{ij} = # occurrences of pattern j in person i

pef_j = # persons containing pattern j

N = # persons in total

EXPERIMENT ON RESEARCHER CLASSIFICATION

Summary of Experiment Dataset

In the experiment, we used the bibliographic datasets of the JST and Elsevier's Scopus (<http://www.scopus.com>). The JST dataset includes 5,000 international journal titles and 9,500 domestic journal titles for science and technology. The Scopus dataset includes 21,000 international journal titles, 417 domestic (Japanese) journal titles. We then retrieved sequential data concerning the following five metrics by year.

- # of papers from domestic journals and conferences
- # of papers from international journals and conferences
- # of citations in domestic papers
- # of citations in international papers
- % of the first author's papers in the total of domestic and international papers

The datasets of researchers include 42 who specialize in Artificial Intelligence (AI) and 72 who specialize in Bioscience (Bio) in Japan (to the best of our knowledge, the order of authors in papers is not alphabetical in those domains). We randomly collected researchers belonging to The Japanese Society for Artificial Intelligence (<http://www.ai-gakkai.or.jp/en>) and The Japanese Society for Regenerative Medicine (<https://www.jsrm.jp/?lang=english>). Then, researchers who had received a grant of more than 30 million JPN as the project representative were labelled as "distinguished" (TRUE in the classification) by referring to the Database of Grants-in-Aid for Scientific Research (<https://kaken.nii.ac.jp/en/>), since grants are given to "distinguished" researchers who propose excellent themes and are determined to achieve those themes through the sufficient deliberation of several domain experts. Sequential data in time were taken for ten years before receipt of the grant, since the grant necessarily increases research achievements. For researchers who do not have such grants (FALSE in the classification), the sequential data were taken in the last decade leading to 2014. The number of researchers with a FALSE classification is much larger than the number of them with TRUE. However, since we assume to have a screening process conducted before the proposed method, the distribution of researchers is set to be equal to each other. Thus, we selected the same number of researchers with and without grants (TRUE and FALSE).

Classification accuracy

This section first presents a baseline result based on the sum of data for ten years concerning the above four metrics and the average for the first author ratio. The features are the numerical values of the achievements instead of their changes in time (k-gram). Table 1 (above) shows the accuracy of researcher classification (TRUE, FALSE) by the 10-fold cross validation using a decision tree, in which the algorithm is C4.5.

Table 1. Classification accuracy (%).

Domain and Features	Class	Precision	Recall	F-measure
AI by quantity	TRUE	80.0	76.2	78.0
	FALSE	77.3	81.0	79.1
Bio by quantity	TRUE	64.2	89.5	74.7
	FALSE	82.6	50.0	62.3
AI by time series and quantity	TRUE	95.0	90.5	92.7
	FALSE	90.9	95.2	93.0
Bio by time series and quantity	TRUE	87.1	75.0	80.6
	FALSE	78.0	88.9	83.1

Next, we combined the feature vectors of the time series patterns representing the changes and the above five features in quantity. The result is shown in Table 1 (below), and thus we can confirm that the combination of both features has higher accuracy than the amount of the achievements alone. We also conducted the Chi-squared test for independence to assess the statistical significance p between the numbers of correctly and incorrectly classified researchers in Tables 1. The results of the AI and Bio domains in Table 1 (below) were superior to those in Table 1 (above) ($p = 0.014$). Thus, there is a statistically significant difference in the classification between the features in quantity and the combination of the time series patterns with them.

RELATED WORK

There are few case studies on time series analysis, and most papers visualize time series changes in terms of specific metrics. For instance, Prathap (2011) proposed exergy as a single number scalar indicator based on a thermodynamic analogy in order to assess the bibliometrics progress of researchers and then represented the progress as a phase diagram. However, Leydesdorff (2013) argues that the sciences evolve as complex and non-linear systems that contain recursive terms and interaction, for example, between universities and industries. Multivariate analysis in bibliometrics has focused mainly on static designs and should address more of its dynamic developments. Bjork, Offer & Soederberg (2014) also proved that there are patterns in the transition of research achievements, as described in the introduction. In terms of the publications of 57 Nobel Prize winners in economics from 1930 to 2005, the study indicated that time series changes in the number of citations received can be classified into four types and also fit an innovation diffusion curve derived from the Bass model. Thus, Kajikawa et al. conducted a Topological Data Analysis of the citation networks of papers. In this study, time series changes in the position of specific papers in the network are represented by three measures of centrality: clustering centrality, closeness centrality, and betweenness centrality; then, the correlation with the number of citations that will be received in the future are estimated (Shibata, Kajikawa & Matsushima 2007), (Iwami et al. 2014). Although the approach is different, the purpose is the same as that of our study.

CONCLUSION AND FUTURE WORK

To provide useful reference information other than the simple sum of the metrics in the examination of research grants, this paper proposed a classification method for researchers based on time series bibliometrics. Future works include an increase in the number of researchers as well as verification in other domains.

REFERENCES

- Bjork, S., Offer, A. & Soederberg, G. (2014). Time series citation data: the nobel prize in economics. *Scientometrics* 98(1):185–196.
- Iwami, S., Mori, J., Sakata, I. & Kajikawa, Y. (2014). Detection method of emerging leading papers using time transition. *Scientometrics* 101(2):1515–1533.
- Leydesdorff, L. (2013). Statistics for the dynamic analysis of scientometric data: the evolution of the sciences in terms of trajectories and regimes. *Scientometrics* 96(3):731–741.
- Patel, P., Keogh, E., Lin, J. & Lonardi, S. (2002). Mining motifs in massive time series databases. In *Proceedings of the 2002 IEEE International Conference on Data Mining (ICDM)*, 370–377.
- Prathap, G. (2011). The energy-exergyentropy (or eee) sequences in bibliometric assessment. *Scientometrics* 87(3):515–524.
- Shibata, N, Kajikawa, Y. & Matsushima, K. (2007). Topological analysis of citation networks to discover the future core articles. *Journal of the American Society for Information Science and Technology* 6(58):872–882.



Inclusion of Gender perspective in scientific publications in Energy Efficiency ¹

Elba Mauleón ^{*}, Daniela De Filippo ^{**}

^{*}mmauleon@bib.uc3m.es; ^{**}dfilippo@bib.uc3m.es

Laboratory for Metric Information Studies (LEMI); Research Institute for Higher Education and Science (INAECU), Carlos III University of Madrid, Calle Madrid, N. 126, Getafe-Madrid, 28038 (Spain)

ABSTRACT

Over the last years, the European Union actively promotes the incorporation of the gender perspective as a transversal category in research and technology, in a way that its presence is considered in all the stages of the scientific process, from the definition of the scientific and technical research priorities, going through research problems, theoretical and explanatory frameworks, the methods, collection and interpretation of data, up to the conclusions, the technical implementations and developments that could arise.

Proof of this is that Horizon 2020 links research and innovation to gender equality in science as a way to reach scientific excellence. It is understood that the scientific method cannot be considered neutral and for this reason men and women should participate on all levels of project research in equal conditions. In fact, gender dimension is included in all the sections of Horizon 2020 to assure that the requirements, singularities, conducts and behaviors of men and women researched are taken into consideration throughout the whole process.

In the field of Energy Efficiency, for example, there are various issues that according to the United Nations affect men and women differently; this is the case of energy efficiency at homes, which is related with the quality of life of people. In developed regions, women have the main responsibility for food production, water supply and energy supply for domestic usage. The women in this regions hold a extra involvement in housework, so that they disproportionately are more exposed from dangerous energy sources such as kerosene that use for instance to maintain the house heated (Khudadad, Sultana, Khan, 2013).

However, many women tend to be underrepresented in the main forums about Energy on all levels. This limits their capacity to contribute and implement their knowledge in the search for solutions and their absence in research on Energy and Energy Efficiency endangers the excellence in this field.

On the basis that quality research should be gender sensitive; the main goal of this work is to analyze the inclusion of gender perspective in the publications about energy efficiency.

¹ This work was supported by the projects: “New horizons in research and innovation for sustainable energy and transport in the urban environment” (REF: CSO2014-51916-C2-1-R) and “Detection of new front of research and innovation in energy efficiency in Spain. Analysis of the flow of knowledge between science, industry and society” (REF: CSO2014-58889-JIN) supported by Ministry of Economy and Competitiveness.

SOURCE AND METHODOLOGY

This study has focused on the analysis of the publications in Web of Science (WoS) related to “Energy Efficiency” in which the content refers to gender issues.

PHASE 1: Identification of documents on Energy Efficiency that include gender dimension

- Selection of documents on Energy Efficiency in WoS through search terms in Title, Keywords and Abstract: TS=("energy efficiency") OR TS=("energy saving") OR TS=("energy storage"). The search was done on three data bases SCI, SSCI, A&HCI without a time limit. The terms used have been extracted from European projects on the subject and have been refined by the use of more ample strategies.
- Detection of the 9 most productive journals on Energy Efficiency.
- Search of content referring to gender in the 10 selected journals: TS="gender" OR TS="female" OR TS="Women" OR TS="Woman".
- Download the documents, filter, treat and standardize the information.
- Elaboration of the bibliometric indicators: number of documents per year, documentary type, language, journal, institution, specialization and country.
- Content analysis (by means of abstract and keywords) to detect the main themes addressed.

PHASE 2: Comparison of documents on Energy Efficiency WITH and WITHOUT the inclusion of the gender dimension.

- Random sampling done to obtain in the selected journals the same number of documents WITHOUT gender dimension
- Collection of the principal bibliometric indicators comparing both subsets: number of authors per document, authorship distribution and the signature order by gender, impact and visibility.

PRELIMINARY RESULTS

The obtained results demonstrate that the most productive journals on Energy Efficiency have been: Energy Power; RSC Advances; Applied Energy; Energy; Electrochemical Acta; Journal of Power Sources; Energy Conversion and Management; Applied Thermal Engineering. In the journals there have been 119 documents detected that included the gender dimension between 1993 and 2016 with a sharp increase in the production since 2010 (Figure 1). The subject categories with which they are associated are: Chemistry; Energy & Fuel; Electrochemistry; Thermodynamics; Environmental Sciences; Construction & Building Technology. Forty-one countries have been responsible for this production, in particular India, USA and China (Figure 2).

The production in these same journals, that do not include gender dimension, spans over a greater period having detected documents dating from 1973. In addition, the distribution by country differs (Figure 2). The impact of these publications WITHOUT gender dimension is greater: 11.85 citations per document opposed to 10.88 citations per document of the documents WITH gender dimension.

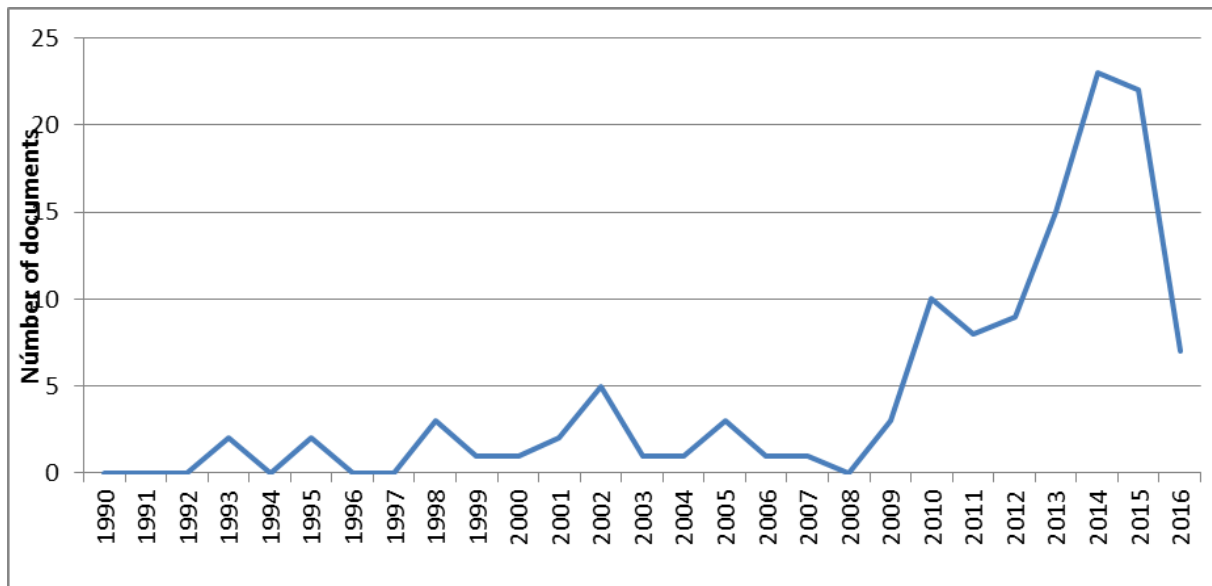


Figure 1. Annual evolution of the number of documents on energy efficiency including the gender dimension

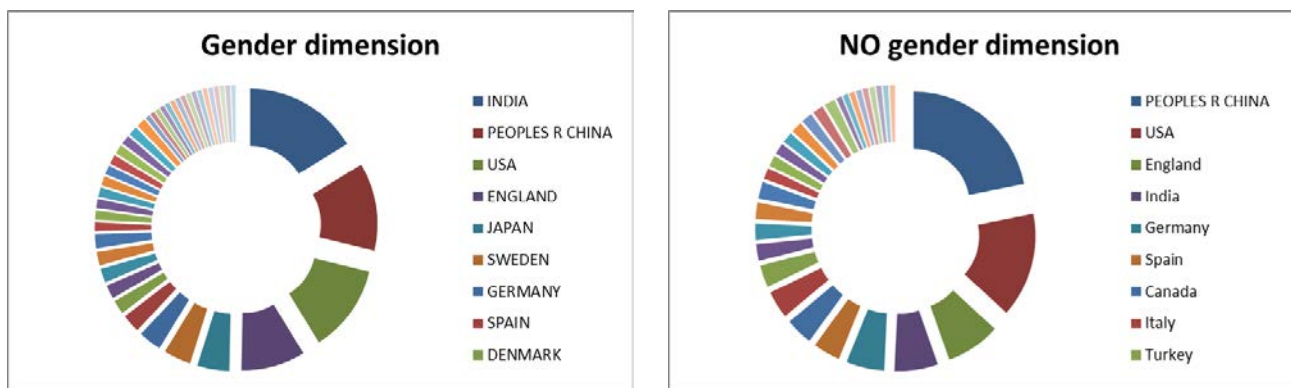


Figure 2. Distribution of of documents by country

CONCLUSIONS

The results obtained show that, despite the importance of a subject such as Energy Efficiency, the inclusion of the gender perspective still seems to be marginal. This is evident in the small number of collected documents that in no journal exceeds 1%. The preliminary results on authorship demonstrate a higher presence of women in the documents that include gender dimension in spite of recommendations about the necessity to attain equality in research groups. The scarce feminine presence in the studied area can be related to the prestige and the extent of consolidation in the included disciplines. Previous studies have shown that when a discipline is emerging, or marginal, the presence of women is more prominent. However, when the disciplines are consolidated and gain prestige, the presence of women begins to become more scarce (Köhler, 1994). In cases like that of Energy Efficiency, in which there is an important presence of relations with Engineering and the development of technology, this orientation could be what creates a bias towards the presence of women. However, nothing would impede the inclusion of the gender perspective in the research content since studies like those done by the European Commission (2013) show the importance of this dimension in priority areas such as energy.

REFERENCES

- European Commission (2013). Gendered Innovation: how gender analysis contributes to research. ISBN 978-92-79-25982-1
- Khudadad, N.; Sultana, N.; Khan, FF. (2013). Women selling and promoting energy efficient products in Northern Pakistan. *Renewable Energy*, 49: 271-274.
- Köhler, R. (1994) Lords of the Fly. Drosophila Genetics and the Experimental Life. Chicago: University of Chicago Press
- United Nations, 2015. http://unfccc.int/gender_and_climate_change/items/7516.php



Gender-based differences in German-language publications

Sabrina J. Mayer*

[*mayer@dzhw.eu](mailto:mayer@dzhw.eu)

German Centre for Higher Education Research and Science Studies (DZHW), Department 2, Schützenstraße 6a, 10119 Berlin (Germany)

ABSTRACT

By now, publishing in another language than English has become a peripheral scientific activity, especially as it is known that non-English publications are generally less cited than English publications. This paper analyses factors that lead to publications in German. From previous research, gender-based differences for scientific output and impact are known. Hence, this paper aims to explore gender-specific patterns in German-language publications. However, to be able to capture the full impact of gender-based differences, three additional factors are considered: 1) the academic field to control for a subject-specific publication culture (with higher shares of female scientists in fields with more German publications), 2) the authorship type (single-authored, domestic, or international collaborations) to control that women are less likely to publish internationally co-written articles, and 3) career length as female researchers are today on average younger/in lower academic positions than male researchers. The preliminary results show that - even when we control for these three additional factors - female-only author teams have a higher probability than male-only author teams to publish in German by 0.6 to seven percentage points.

INTRODUCTION

English is the current scientific lingua franca; especially in countries with a small scientific community publishing in English enables the researchers to integrate and communicate with other researchers. For an analysis of publication patterns, German is chosen because it is still the second most common language (0.72 percent) in Web of Science (WoS). However, the total number of publications in German declined by 13,686 from 1983 to 2013. By now, publishing in German has become a peripheral scientific activity. So what drives publishing in German? It is known that there are gender-based differences in publication output and citation impact, although these results are sometimes inconsistent because often only specific subfields are analyzed (e.g. Larivière, Ni, Gingras, Cronin, & Sugimoto, 2013; Maliniak, Powers, & Walter, 2013; Paul-Hus et al., 2015; West, Jacquet, King, Correll, & Bergstrom, 2013). We know from research on Russia that women are more likely to publish in the national language than men (Paul-Hus, Bouvier, & Ni et al., 2015). However, these results were obtained without controlling for the discipline/field.

This paper aims to explore gender-specific patterns in own-language publications. However, analyzing the relationship between gender and publication language without additional factors falls short to capture the genuine impact of gender so three other factors are considered.

First, the academic field of a publication needs to be controlled for. Own-language publications are more common in fields that are often the fields where higher shares of female scientists are found (She Figures 2012, 2013). An analysis of own-language publications that does not include the academic field may find gender-based differences that are simply caused by the different subject-specific publication cultures.

Second, the type of authorship needs to be considered: Previous studies found gender-specific patterns with regard to collaborations: men rather publish in international co-authorship whereas women rather collaborate nationally (Fisher, Cobane, Vander Ven, & Cullen, 1998; Maliniak, Powers, & Walter, 2013). It can be expected that internationally co-written papers are less likely written in German.

Third, the current career position needs to be accounted for. In recent years, more and more women entered academia, although these shares vary by discipline (e.g. Burelli, 2008). Hence, the age structure/academic ranks of women and men are different: for women, the share of scientists with entry-level positions is higher. Differences in own-language publications could be caused by this difference in career length: The longer somebody they are part of the academia, the more likely non-native speakers are to know the rules of the game to successful publications in high-ranking English-language journals.

DATA BASE AND METHODS

An in-house data base of Web of Science is used. Publications in German are only likely if somebody actually has a high level of German, so only publications with at least one author with a German address are included. To capture the current point of developments in German-language publication, 2013 is chosen.

The academic field is controlled for by using the OECD fields of knowledge for which a mapping is provided by Thomson Reuters, Humanities is excluded because of the still persisting coverage gap.

So far, no established approach for gender identification exists; often, matching of first names (in some cases, such as for Russian names, by last names) to name lists is done (Larivière, Ni, & Gingras et al., 2013; Maliniak, Powers, & Walter, 2013; Paul-Hus, Bouvier, & Ni et al., 2015; West, Jacquet, & King et al., 2013). Most existing studies use the data set from the US Social Security Administration, that contains the 1,000 most common female and male first names by year, available in the R package “gender”. Because some first names are androgynous or depend on regional origin, first names are only classified when they were given in 1970 in 95 percent of cases to a single gender. However, as the dataset originates from the United States, the package “gender.c”¹ is also used (see Table 1).

Academic career length will be approximated by the first publication available in WoS². This approach is biased because a change of name after marriage is much more likely for women, so this approach might underestimate the career length for women. In addition, a distinct author disambiguation is still missing for WoS, so authors with the same first- and last name cannot be clearly distinguished.

¹ Michael, Jörg (2007). 40 000 Namen. In: c't n. 17: 182–183. <ftp://ftp.heise.de/pub/ct/listings/0717-182.zip>

² Previous studies showed a high correlation between the length of an academic career and the first publication, depending on the academic field (Costas, Nane, & Lariviere, 2015).

Table 1: Gender identification by discipline

	Number of papers in total	Number of authors in total	Authors with first name only initials	Analyzed papers/ authors	Gender Identification (percentages)
Agricultural Sciences	7,039	12,536	7,428	1,729 5,108	38.2 f 58.1 m 3.7 NA
Engineering and Technology	30,127	88,191	32,426	16,345 55,765	20.5 f 72.8 m 6.7 NA
Medical and Health Sciences	52,858	219,941	82,019	33,401 137,922	29.5 f 65.8 m 4.7% NA
Natural Sciences	67,948	296,756	114,367	52,069 182,389	24.6 f 69.5 m 5.9 NA
Social Sciences	11,503	24,436	2,309	10,392 22,127	31.3 f 63.8 m 4.9 NA

100 authors from each field are randomly drawn and manually checked for gender and actual career length to know about the error margin and possible gender-differences of the error.

RESULTS

On average, papers written by only women are twice as likely to be written in German than those written by all men teams (7.5 to 2.4 percent).

Logistic regression is performed with German-language publication as dependent variable, including all relevant predictors (see Table 2).

Table 2: Logistic Regression on the Probability to publish in German, logits

	DV: Publication in German
Academic Field (Ref.Group: Agricultural Sciences)	
Engineering and Technology	-1.574*** (0.073)
Medical and Health Sciences	-1.523*** (0.069)
Natural Sciences	-3.798*** (0.084)
Social Sciences	-1.765*** (0.075)
No. Of Authors	-0.356*** (0.011)
No. Of Countries	-1.470*** (0.050)
Average Career Length	-0.043*** (0.008)
Gender (Ref. Group: Female)	
Male	-0.249*** (0.051)
Mixed-gender author teams	-0.579*** (0.060)
Constant	2.609*** (0.099)
Observations	112,209
Log Likelihood	-13,452.920
Akaike Inf. Crit.	26,925.840
McFadden's Pseudo R ²	0.25

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

When all other factors are controlled, female-only publications are still more likely to be in German, e.g. publications with two authors from Germany in Medical and Health Sciences (Natural Sciences) with an average career length of 3 years, have a predicted probability of 22.7 (2.9) percent to be in German for female-only author teams and of 18.6 (2.3) percent for male-only author teams.

DISCUSSION

My results show that there are differences in publication behaviour of women and men concerning the chosen publication language. These differences diminish when other predictors, such as academic field, career length, and collaboration type are taken into account, but do not vanish. This could probably be caused by the rather broad academic fields used, or by the career length measure that is only a rough proxy. However, I could show that all relevant variables need to be included because the gender-based differences decrease when other predictors were considered, so not including them would draw a wrong picture of the linkages. Still, one could wonder if motivational structures and “academic self-esteem” cause this; one could suppose that men are more likely to try to publish in higher ranking journals and tend to build more international networks, but these indicators are not available for bibliometric analyses.

REFERENCES

- Burelli, J. (2008). *Thirty-Three Years of Women in S&E Faculty Positions*. National Science Foundation.
- Costas, R., Nane, T., & Larivière, V. (2015). Is the Year of First Publication a Good Proxy of Scholars' Academic Age? *Proceedings of the 15th International Conference of the International Society for Scientometrics and Informetrics*, 988–998.
- Fisher, B. S., Cobane, C. T., Vander Ven, T. M., & Cullen, F. T. (1998). How Many Authors Does It Take to Publish an Article? Trends and Patterns in Political Science. *PS: Political Science and Politics*, 31(4), 847–856.
- Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature*, 504(7479), 211–213.
- Maliniak, D., Powers, R., & Walter, B. F. (2013). The Gender Citation Gap in International Relations. *International Organization*, 67(04), 889–922.
- Paul-Hus, A., Bouvier, R. L., Ni, C., Sugimoto, C. R., Pisljakov, V., & Larivière, V. (2015). Forty years of gender disparities in Russian science: A historical bibliometric analysis. *Scientometrics*, 102(2), 1541–1553.
- She Figures 2012 (2013). *Gender in Research and Innovation. EUR: Vol. 25617*. Luxembourg: Publ. Off. of the Europ. Union.
- West, J. D., Jacquet, J., King, M. M., Correll, S. J., & Bergstrom, C. T. (2013). The role of gender in scholarly authorship. *PloS one*, 8(7), e66212.



Scientific productivity and the impact of neurosurgery scientists in WOS and Mendeley: a gender study

Hajar Sotudeh^{*}, Tahereh Dehdarirad^{**} and Aida Pooladian^{**}

^{*} sotudeh@shirazu.ac.ir

Department of Knowledge and Information Sciences, Faculty of Education and Psychology, Shiraz University, Eram Campus, 71946-84471, Shiraz (Iran)

^{**} tdehdari@gmail.com; ^{**} aida.pouladian@gmail.com

Department of Library and Information Science, University of Barcelona, Melcior de Palau, 140, 08014 Barcelona (Spain)

ABSTRACT

This study seeks to investigate scientific performance of female and male authors in the field of neurosurgery in terms of their frequency, scientific productivity, as well as their impacts (their citation and readership counts). A total of 18851 articles published during 2008-2012 in the eight neurosurgery core journals were retrieved from the Thomson Reuters Web of Science database. Only 1728 (9.16 %) of the papers were found to have readership in Mendeley, comprising 3099 unique authors. Descriptive and inferential (Chi square, T-test) statistics were applied to analyse the data. The results suggested a significant gender difference in terms of both frequency of authors and the number of papers in favour of male authors. Regarding the scientific impact, the results showed no gender differences in terms of average citations received per paper. However, male authors showed a significant superiority in terms of their average readership per paper.

INTRODUCTION

Science has recently witnessed greater participation and contribution from women (Fox, 2005). Although, some research confirms gender equality regarding their number of scientific papers or citations (Van Arensbergen et al. 2012; Copenheaver et al., 2010; Slyder, et al, 2011; Sotudeh and Khoshian, 2014), gender inequality has not yet been fully eradicated. It is, therefore, crucial to constantly monitor females' scientific performance, to understand the underlying factors and thereby eradicate the inequalities (Fox, 2005).

STATEMENT OF THE PROBLEM

Women play an active role in cyberspace, especially in social networks (Kimbrough et al., 2013; Duggan et al., 2015). Given the democratic nature of cyberspace, it is expected that social networking provides a sexually-neutral milieu, where both genders benefit from the increase in their recognition regarding web citations (Kretschmer and Aguillo, 2005), social event counts (Paul-Hus et al., 2015), citations (Sotudeh and Khoshian, 2014) and profile views (Thelwall and Kousha, 2015). To test the degree of gender-neutrality of social media, it is necessary to conduct further gender comparisons in scholarly communication and social networks (Paul-Hus et al., 2015).

Choosing Neurosurgery field as a sample, the present study investigates genders' frequency, scientific productivity, as well as impacts -including citation and readership counts.

METHODOLOGY

Data collection

Using WoS, we downloaded bibliographical data of 18851 articles published during 2008-2012 in the eight Neurosurgery core journals, identified by Venable et al. (2016). We used Webometric Analyst to extract readership counts via Mendeley API. After standardization of names, a list of 3099 unique authors was identified including 2419 (78.1%) males and 680 (21.9%) females.

Data analysis

Due to the skewness of citation and readership distributions (Thelwall and Sud, 2015), bootstrapping techniques were applied to calculate a 95% confidence interval (CI) for the means of both scientific impact measures.

RESULTS

Genders' frequency

The results of Chi square showed that the number of females was significantly lower compared to their male counterparts ($\chi^2=975.83$, $P<0.001$).

Comparing the genders' scientific productivity and impact

The results of T-test showed a significant difference between the genders regarding mean number of their papers, in favor of men ($t=3.01$, $P=0.003$) (Table 1).

Table1. T-test for comparing the genders' mean productivity

Gender	No.	Mean (std)	t	Sig.
female	680	1.07 (0.32)	3.01	0.003
male	2419	1.12 (0.43)		

The comparison of genders regarding their citations per paper showed no significant differences ($t=1.33$, $P=.14$). However, the males exhibited a significant superiority ($t=2.49$, $P=0.007$) in their average readership per paper ($M=3.07$) compared to the females ($M=2.65$) (Table 2).

Table 2. Bootstrapping t-test for comparing the genders' impacts

Measure	Mean (CI)		Mean differences (Female-male)	t	Sig (two-tailed)
	Female	Male			
Citation	74.08 (60.22-88.90)	87.18 (77.49-96.08)	-13.1	1.33	0.14
Readership	2.65 (2.42-2.92)	3.07 (2.92-3.23)	-0.42	2.49	0.007

CONCLUSION AND DISCUSSION

Our study investigated genders' scientific performance in neurosurgery in terms of their frequency, scientific productivity, and impacts (citation and readership counts). Although, our results suggested a significant male superiority in their frequency and their number of papers, we found no evidence of gender bias in terms of average citations received per paper. The impact equality is in line with some previous studies (Copenheaver et al., 2010; Sotudeh and Khoshian, 2014), while contradicting some others (Lewison and Markusova, 2011; Pudovkin et al., 2012). As indicated by Østby et al. (2013) the evidence is inconclusive with regard to citation impact of female and male researchers. Thus, one cannot simply derive a general pattern in terms of the sexuality impact on citations and further studies are needed to corroborate these conclusions.

The males showed a significant superiority in their average readership per paper. This might be due a lower presence of women in Mendeley, or gender differences in choosing social media type to communicate their scholarly findings. There is growing evidence that men and women use online social platforms differently (Volkovich, et al., 2014).

REFERENCES

- Copenheaver, C. A., Goldbeck, K., & Cherubini, P. (2010). Lack of gender bias in citation rates of publications by dendrochronologists: what is unique about this discipline?. *Tree-Ring Research*, 66(2), 127-133.
- Duggan, M., Ellison, N. B., Lampe, C., Lenhart, A., & Madden, M. (2015). Social Media Update 2014. Pew Research Center's Internet & American Life Project. Retrieved March 11, 2016, from <http://www.pewinternet.org/2015/01/09/social-media-update-2014/>
- Fox, M. F. (2005). Gender, Family Characteristics, and Publication Productivity among Scientists. *Social Studies of Science*, 35(1), 131-150. doi: 10.1177/0306312705046630
- Kimbrough, A. M., Guadagno, R. E., Muscanell, N. L., & Dill, J. (2013). Gender differences in mediated communication: Women connect more than do men. *Computers in Human Behavior*, 29(3), 896-900.
- Kretschmer, H., & Aguillo, I. F. (2005). New indicators for gender studies in Web networks. *Information Processing and Management*, 41, 1481-1494.
- Lewison, G., & Markusova, V. (2011). Female researchers in Russia: Have they become more visible? *Scientometrics*, 89, 139-152.

- Østby, G., Strand, H., Nordås, R., & Gleditsch, N. P. (2013). Gender gap or gender bias in peace research? Publication patterns and citation rates for *Journal of Peace Research*, 1983–2008. *International Studies Perspectives*, 14(4), 493-506.
- Paul-Hus, A., Sugimoto, C. R., Haustein, S., & Larivière, V. (2015). Is there a gender gap in social media metrics?. In *Procs of ISSI 2015-15th International conference of the International Society for Scientometrics and Informetrics* (pp. 37-45).
- Pudovkin, A., Kretschmer, H., Stegmann, J., & Garfield, E. (2012). Research evaluation. Part I: productivity and citedness of a German medical research institution. *Scientometrics*, 93(1), 3-16.
- Slyder, J. B., Stein, B. R., Sams, B. S., Walker, D. M., Beale, B. J., et al. (2011). Citation pattern and lifespan: A comparison of discipline, institution, and individual. *Scientometrics*, 89, 955–966.
- Sotudeh, H., & Khoshian, N. (2014). Gender, web presence and scientific productivity in nanoscience and nanotechnology. *Scientometrics*, 99(3), 717-736. doi: 10.1007/s11192-014-1234-6
- Thelwall, M., & Kousha, K. (2015). ResearchGate: Disseminating, communicating, and measuring Scholarship? *Journal of the Association for Information Science and Technology*, 66(5), 876-889. doi: 10.1002/asi.23236
- Thelwall, M., & Sud, P. (2015). Mendeley readership counts: An investigation of temporal and disciplinary differences. *Journal of the Association for Information Science and Technology*, n/a-n/a. doi: 10.1002/asi.23559
- Van Arensbergen, P., van der Weijden, I., & van den Besselaar, P. (2012). Gender differences in scientific productivity: a persisting phenomenon? *Scientometrics*, 93(3), 857-868. doi: 10.1007/s11192-012-0712-y
- Venable, G. T., Shepherd, B. A., Loftis, C. M., McClatchy, S. G., Roberts, M. L., Fillinger, M. E., Klimo, P., Jr. (2016). Bradford's law: identification of the core journals for Neurosurgery and its subspecialties. *J Neurosurg*, 124(2), 569-579. doi: 10.3171/2015.3.jns15149
- Volkovich, Y., Laniado, D., Kappler, K. E., & Kaltenbrunner, A. (2014). Gender patterns in a large online social network. In *International Conference on Social Informatics* (pp. 139-150). Springer International Publishing.



How is the counting method for a publication or citation indicator chosen?

Marianne Gauffriau

Faculty Library of Natural and Health Sciences, Copenhagen University Library, The Royal Library, Nørre Alle 49, DK-2200 Copenhagen N, Denmark, mgau@kb.dk,

SUND Research & Innovation, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, DK-2200 Copenhagen N, Denmark, marianne.gauffriau@sund.ku.dk

ABSTRACT

There is no consensus on when to use a specific counting method (CM) to assign publication and citation scores to an object of study in a bibliometric analysis. This study presents an analysis of the choice of CM in 11 recent studies from Journal of Informetrics: Ten studies use whole scores, and one study use fractional scores. Four studies argue for the choice of CM.

The sample is too small to draw firm conclusions, but a preliminary categorization of types of arguments was made. The categories are: Attributes of authorship, Simplification, Availability, Prevalence, Additive, and Conclusion of study is unaffected. The category Attribute of authorship was identified in three of the four studies. Only the sub-category, Contribution, was linked to more CMs. Otherwise, there was a one-to-one relation between sub-categories and CMs. This may indicate that different CMs are perceived as fit for different purposes.

INTRODUCTION

This study investigates if and how recent bibliometric studies discuss the choice of counting method (CM) when assigning publication and citation scores to objects of study, e.g. countries. Ultimately, arguments for CMs may provide new perspectives on what publication and citation indicators measure.

A CM can affect the result of a bibliometric analysis. For countries with a high degree of international co-authorship whole scores will result in inflated publication indicators (Gauffriau *et al.*, 2008). Similar effects of collaboration is seen when calculating field-normalizations (Waltman & van Eck, 2015). The latter study discusses arguments for and against whole and fractional scores: Which is the more intuitive, the two CMs measure participation and contribution, respectively, and how contribution and collaborations in general are measured (*ibid.* p.889-890). There is no consensus on when to use a specific CM.

METHOD

I analyzed if and how recent bibliometric studies discussed the choice of CM. Larsen (2008) did a similar study and concluded that only five of 85 studies from the International Society for Scientometrics and Informetrics's proceedings discussed the choice of CM (*ibid.*, p.237). Larsen focused on whether the best choices were made. My focus was on the arguments for CMs.

Of the 26 studies in the first 2016 issue Journal of Informetrics, I analyzed the 14 studies which included publication and citation indicators in the method and result sections. Studies on how to handle fields in field-normalization, network analysis, altmetrics etc. were excluded. For many of these, CMs is indeed relevant to discuss, but the argumentation may follow a different logic than arguments for CMs for publication or citation indicators.

The analysis is a snapshot of the state of art and do by no means cover all CMs or arguments for CMs.

RESULT

The choice of CM is indifferent in three studies which measure citation scores for papers, journals and single-authors (Letchford *et al.*, 2016; Haddawy *et al.*, 2016; Bouyssou & Marchant, 2016). There is only one possible value of an object of study (e.g. one journal title) per publication. Therefore a change of CM does not have any effect.

Of the remaining 11 studies, ten use whole scores, and one uses fractional scores. Four studies explicitly discuss the choice of CM. In Table 1, I attempt to assign the arguments to general categories.

Table 1. Arguments for CMs.

Argument	... for CM*	Argument category
“...full counting method can thus be seen as measuring participation...” (Cimini <i>et al.</i> , 2016, p.204)	Whole	Attributes of authorship (participation)
“...fractional counting as measuring how many papers are creditable to a country...” (ibid.)	Fractional	Attributes of authorship (volume)
“...reasonable to accept the simplification that all authors’ contributions are equally important. [...] ...however, the different types of collaboration should be considered as possible explanations for any patterns found.” (Thelwall & Sud, 2016, p.49)	Whole	Attributes of authorship (contribution)
“...fractional contribution of researcher to publication...” (Abramo <i>et al.</i> , 2016, p.34)	Fractional	
“...authors in simple alphabetical order: in this case the fractional contribution simply equals the inverse of the number of authors.” (ibid.)	Fractional, rank-independent	
“...contributions to the published research by the order of the names in the listing of the authors. [...] ...network based flexible allocation scheme, giving different weights to each co-author according to their position in the list of authors and the character of the co-authorship (intra-mural or extra-mural)” (ibid.)	Fractional, rank-dependent	
“...the data contains no overlap of individual networks of different reference institutions or at least it was possible	Whole	Simplification

to ignore this overlap; [...] Multiple co-authorships with authors from different institutions create duplicates of papers [overlap]... [...] Measurement dependencies [overlap] are taken into account to some extent by a multilevel statistical modelling strategy.” (Bornmann <i>et al.</i> , 2016, p.316)		
“...fractional counting is complicated by the several ways in which weights can be assigned...” (Cimini <i>et al.</i> , 2016, p.204)	Whole	
“...full counting is adopted as the SCImago statistics are built according to this criterion.” (<i>ibid.</i>)	Whole	Availability
“...full counting is also commonly adopted...” (<i>ibid.</i>)	Whole	Prevalence
“...fractional counting [...] leads to a proper field normalization of impact indicators...” (<i>ibid.</i>)	Fractional	Additive (avoid double scores)
“...the difference between the two methods [whole and fractional] basically consists in a country-specific rescaling of impact indicators, the relative temporal changes of countries impact we will analyse are, per se, unaffected by the choice of the counting method.” (<i>ibid.</i>)	Whole	Conclusion of study is unaffected

* Terminology: Gauffriau *et al.* (2007, p.178-180) except that fractional replaces normalized.

CONCLUSION

Among the 11 studies, the most common CM is whole scores. Seven studies do not argue for the choice of CM. In the four remaining studies, the category Attribute of authorship is identified in three studies. Only the sub-category, Contribution, is linked to more CMs. Otherwise, there is a one-to-one relation between sub-categories and CMs. This may indicate that different CMs are perceived as fit for different purposes.

REFERENCES

- Abramo, G. *et al.* (2016). A methodology to measure the effectiveness of academic recruitment and turnover. *Journal of Informetrics*, 10(1), 31–42.
- Bornmann, L. *et al.* (2016). Excellence networks in science: A Web-based application based on Bayesian multilevel logistic regression (BMLR) for the identification of institutions collaborating successfully. *Journal of Informetrics*, 10(1), 312–327.
- Bouyssou, D., & Marchant, T. (2016). Ranking authors using fractional counting of citations : An axiomatic approach & . *Journal of Informetrics*, 10(1), 183–199.
- Cimini, G. *et al.* (2016). Investigating the interplay between fundamentals of national research systems: Performance, investments and international collaborations. *Journal of Informetrics*, 10(1), 200–211.
- Gauffriau, M. *et al.* (2007). Publication, cooperation and productivity measures in scientific research. *Scientometrics*, 73(2), 175–214.

- Gauffriau, M. *et al.* (2008). Comparisons of results of publication counting using different methods. *Scientometrics*, 77(1), 147–176.
- Haddawy, P. *et al.* (2016). A comprehensive examination of the relation of three citation-based journal metrics to expert judgment of journal quality. *Journal of Informetrics*, 10(1), 162–173.
- Larsen, P. O. (2008). The state of the art in publication counting. *Scientometrics*, 77(2), 235–251.
- Letchford, A. *et al.* (2016). The advantage of simple paper abstracts. *Journal of Informetrics*, 10(1), 1–8.
- Thelwall, M., & Sud, P. (2016). National, disciplinary and temporal variations in the extent to which articles with more authors have more impact: Evidence from a geometric field normalised citation indicator. *Journal of Informetrics*, 10(1), 48–61.
- Waltman, L., & van Eck, N. J. (2015). Field-normalized citation impact indicators and the choice of an appropriate counting method. *Journal of Informetrics*, 9(4), 872–894.



The occurrence areas of the dependence problem of the h-index¹

Chichen Liu^{*}, Sanfa Cai^{**} and Yuxian Liu^{***}

^{*} 1531071@tongji.edu.cn

School of Economics and Management, Tongji University, 4800 Cao An Road 4800, Shanghai, 201804 (China)

^{**} csf@tongji.edu.cn

School of Economics and Management, Tongji University, 1239 Siping Street, Shanghai, 200092 (China)

yxliu@tongji.edu.cn;

Tongji University Library, Tongji University, 1239 Siping Street, Shanghai, 200092 (China)

ABSTRACT

The aim of this paper is to figure out the situations that will arise the problem of dependency of the h-index. In doing so we start from the study of some specific examples to find out the contributing factors to this problem. With the help of mathematical derivation, we figure out three situations that will not cause the problem of dependency. Then we introduce the coordinate systems to help us divide the left possible situations into five areas. And we use the enumeration method to check the five areas in very single coordinate systems to further determine the situations that will arise the problem of dependency. We conclude that as long as one of the four situations is satisfied: (1) the number of the new papers is not larger than the difference of the h-indices of the two authors; (2) the numbers of the citations received by the new papers are all not larger than the larger h-index of the two authors; (3) the number of the new papers and the numbers of citations received by X new paper are all larger than the maximum of the maximum of the citations of the original h-core papers written by the two authors; (4) the maximum of the citations of the original papers written by the author with lower h-index is not larger than the original h-index of the author with higher h-index, the problem of dependency of the h-index will absolutely not occur. And the other situations excluding the four situations above could arise the problem of dependency but the probability is not 100 per cent.

INTRODUCTION

When ranking scientists or journals or department, they defined a series of axioms such as nontriviality, independence, archimedeaness, transferability, consistency, homogeneity, transfer, monotonicity, and dummy paper, weak order. As a result, scientists, journals and departments can be ranked consistently (Marchant, 2009; Bouyssou & Marchant, 2010; 2011; 2014). Among these axioms, independence is defined as follow: *If we rank two scientists L and S according to an indicator, it satisfies that*

For all $L, S \in A$, all $x \in \mathbb{N}$, $L \gtrsim S$ iff $L + \sum_{l=1}^x I_{x_l} < S + \sum_{l=1}^x I_{x_l}$,

Then we say this indicator is dependency.

Otherwise For all $L, S \in A$, all $x \in \mathbb{N}$, $L \gtrsim S$ iff $L + \sum_{l=1}^x I_{x_l} \gtrsim S + \sum_{l=1}^x I_{x_l}$,

We say this indicator satisfies independency.

Here I_{x_l} denotes that the paper ranked in the l^{th} position receive x_l citations.

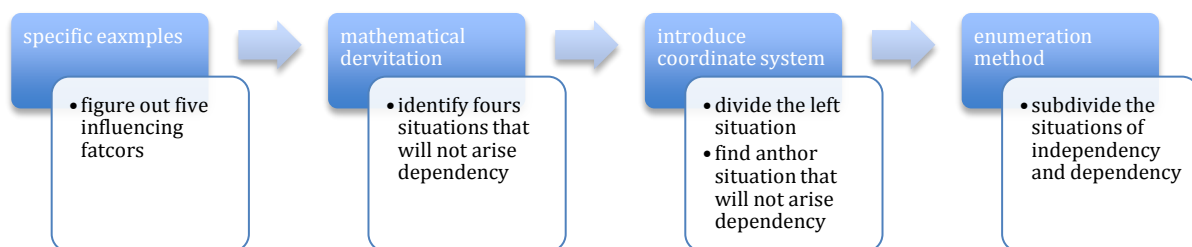
¹ This work was supported by NSFC via 71173154, National Social Sciences Foundation of China (08BZX076) and the Fundamental Research Funds for the Central Universities.

Many scholars have realized that many indicators have the problem of dependency. Especially the dependency of the h-index has been discussed (Prathap, 2012; Waltman, L & van Eck, 2012). Liu (2013) calculated the probability of dependency. We take the initiative to analyze what factors contribute to the problem of dependency, what situation will arise the problem of dependency and what is the probability of dependency.

METHODS

We use the h-index to illustrate our ideas. The h-index was decided by the multiset of the number of the citations the publication in the h-core have received. From the mathematical property of this multiset we figured out five factors that contribute to the occurrence of dependency of the h-index, four situations that will not (and the other areas that will) cause the problem of dependency. (Figure 1 is the diagram showing our roadmap)

Figure 1, roadmap diagram



RESULTS

Five factors that influence the occurrence of the dependence

The h-index is not always independence or dependence. There are five factors such as the original h-index of author L (h_L), the original h-index of author S (h_S), the maximum number of the citations of the original h-core papers of author L (m_L), the maximum number of the citations of the original h-core papers of author S (m_S), the maximum of m_L and m_S (M) can influence the occurrence of the dependence

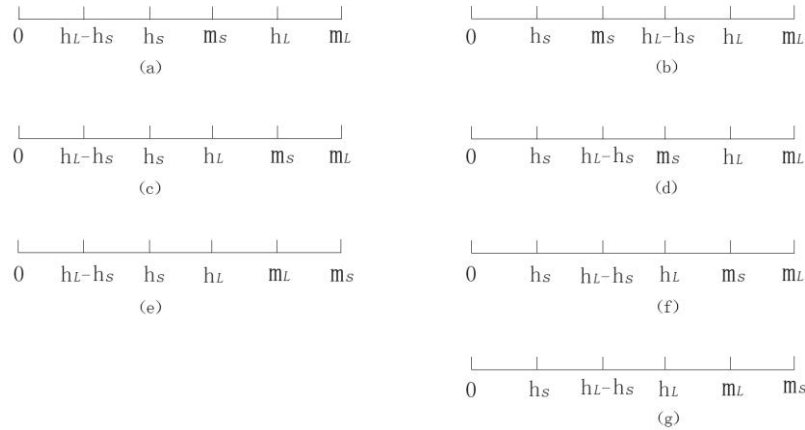
Four Situations that will not arise dependency

Through mathematical derivation, we conclude that as long as one of the four situations is satisfied: (1) the number of the new papers is not larger than the difference of the h-indices of the two authors; (2) the numbers of the citations received by the new papers are all not larger than the larger h-index of the two authors; (3) the number of the new papers and the numbers of citations received by X new paper are all larger than the maximum of the maximum of the citations of the original h-core papers written by the two authors; (4) the maximum of the citations of the original papers written by the author with lower h-index is not larger than the original h-index of the author with higher h-index, the problem of dependency of the h-index will absolutely not occur.

Subdivision of the left situation

Then we introduce the axis (figure 2) to show the possible size relationship between the five contributing factors.

Figure 2: the possible size relationship between the five contributing factors

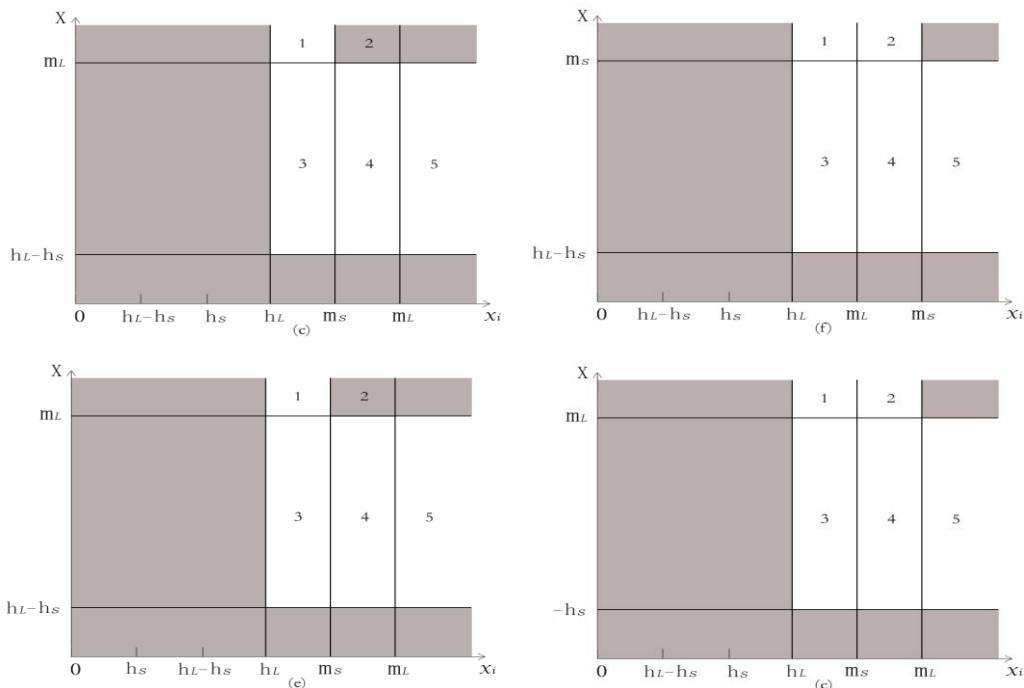


According to the size relationship in figure 2 where there are actually four situations (c, e, f, g) that will arise the problem of dependency because the situation $h_L \geq m_S$ will not arise dependency.

We assume author L and author S publish X articles with x_i citations.

we build up four corresponding coordinate systems in figure 3. we use the enumeration methods to check whether the problems of occurrence of dependence can happen in these areas or not. The shadows show the situation of independency, that means will not arise the problem of dependency. The white part represents uncertain situations where both independency and dependency of h-index could happen.

Figure 3: situations that will and will not arise the problem of dependency



In the white areas the new h-index h'_L could be any integer in the interval $[h_L, h_L + X]$, and h'_S could be any integer in the interval $[h_S, h_S + X]$. Only $h'_S > h'_L$ can make the problem of dependency occur, which will happen in the grey grids in Figure 4.

Figure 4: possible h-indices of the author L and S after publishing X new papers

$h_S' \backslash h_L'$	h_L	h_L+1	h_L+2	...	h_S+X-1	h_S+X	h_S+X+1	...	h_L+X
h_S									
h_S+1									
h_S+2									
...									
h_L									
h_L+1									
h_L+2									
...									
h_S+X-1									
h_S+X									

CONCLUSION

In this article, we find the factors that influence the occurrence of dependency of the h-index. We find there exist four situations that will not arise the problem of dependency of the h-index: (1) $X \leq h_L - h_S$; (2) $x_i \leq h_L$; (3) $X \geq M$ and $x_i \geq M$ and (4) $m_S \leq h_L$. Apart from these four situations, the left are all what could arise the problem of dependency but the probability is not 100%. In our future study, we will try to further subdivide the situations of dependency and calculate the possibility.

REFERENCES

Bouyssou, D., & Marchant, T. (2010). Consistent bibliometric rankings of authors and journals. *Journal of Informetrics*, 4, 365–378.

Bouyssou, D., & Marchant, T. (2011). Ranking authors and departments in a consistent manner. *Journal of the American Society for Information Science and Technology*, 62(9), 1761–1769.

Bouyssou, D., & Marchant, T. (2014). An axiomatic approach to bibliometric rankings and indices. *Journal of Informetrics*, 8, 449–477.

Prathap, G. (2012). The inconsistency of the H-index. *Journal of the American Society for Information Science and Technology*. 63 (2), 1480-1482.

Marchant, T. (2009). An axiomatic characterization of the ranking based on the h-index and some other bibliometric rankings of authors. *Scientometrics*, 80(2), 325–342.

Waltman, L., & van Eck, N.J. (2012). The inconsistency of the h-index. *Journal of the American Society for Information Science and Technology*, 63(2):406-415

Liu, Y. (2013). To what extent is the h-index inconsistent? Is strict consistency a reasonable requirement for a scientometric indicator? *Proceedings of ISSI 2013 Vienna*, (Juan Gorraiz, Edgar Schiebel, Christian Gumpenberger, Marianne Hörlesberger, Henk Moed, eds.), AIT Austrian Institute of Technology, Vienna, p. 1696-1710.



Multivariate bibliometric analysis of scientific production indicators: a taxonomy of countries scientific degree of centrality

Deise Deolindo Silva* and Maria Cláudia Cabrini Grácio**

* deisedeolindo@hotmail.com

PhD student of post graduate degree in Information Science at UNESP – Univ Estadual Paulista and teacher of Fatec-Garça. Address: 2331, Presidente Vargas Avenue, Garça, 17400-000, (Brazil).

** cabrini@marilia.unesp.br

UNESP – Univ Estadual Paulista, 737 Hygino Muzzi Filho Avenue, Marília, 17525-900 (Brazil).

In recent decades, there has been a significant growth in the number of electronic scientific publications. Socialization of knowledge can be represented by a spiral cycle of scientific production and communication, in which knowledge is the effect of social relations, from a social consensus, so it tends to be ascendant (Guimarães, 2000).

With the increase of scientific publications, the evaluation of this production has become an essential factor that allows reflection on the degree of development of a country and provides social, political and economic guidelines.

In this context, it is necessary to develop indicators to reliably evaluate the various scientific areas and participation of countries, contributing to a more realistic view of the world scientific behavior with subsequent definitions of political and scientific guidelines, allocation of investments and resources, programs and activities related to scientific and technological development (Mugnaini *et al.*, 2004; Thomaz *et al.*, 2011; Vanz & Stumpf, 2010).

Among the approaches to characterize and evaluate a scientific area, bibliometric studies stand out, as they have tangible and reliable procedures (Oliveira & Grácio, 2011). Among the bibliometric indicators of impact, are: h-index; number of produced documents; number of citations and citations per document. As infrastructure indicator of countries, Human Development Index (HDI) is considered, which according to PNUD (2016), is a concise measure of long-term progress, which considers three basic dimensions: income, education and health.

In this context, the joint analysis of bibliometric indicators, through multivariate analysis, contributes to a better understanding and visualization of the world scientific behavior.

Among the multivariate analysis, the cluster analysis stand out, as it allows a taxonomy of individuals (Hair *et al.*, 2009), whether they are authors or countries participating in the produced science to explain the degree of similarity and dissimilarity among these individuals, from the combined analysis of multiple variables.

Considering the above, this research aims to analyze the scientific production of countries participating in the mainstream science in the 1996-2014 period. More specifically, the objective is to group the countries into similar groups, designated as central, emergent and peripheral countries, depending on the joint analysis of their scientific performance in relation

This work is licensed under a Creative Commons License: Attribution-NonCommercial-NoDerivatives 4.0 International

to indicators: h-index; number of produced documents; number of citations and citations per document. Moreover, we aim to compare the ranking of countries in relation to h-index, with HDI value.

As a research procedure, the research universe was the scientific production of 194 countries in the period of 18 years (1996-2014), with h-index of at least 18. We obtained on *SCImago Journal & Country Rank*, in February 2016, for each country, the h-index, the number of documents, citations and citations per document. We sought the HDI value of the countries in PNUD (2015). We used the cluster analysis technique using k-means to obtain the desired clusters. The Wilcoxon-Mann-Whitney test was used to compare group means for each indicator.

The Central group is composed by 9 countries, Emergent by 24, and Peripheral by 161 nations. The Central cluster has the highest values for the indicators and higher values for the HDI. The Emergent countries have median production and high HDI, the Peripherals have lower values for the bibliometric indicators and median HDI.

Table 1 shows the statistical summary of the bibliometric indicators for the clusters.

Table 1: statistical values of groups

<i>Índice h</i>				
<i>Cluster</i>	<i>Minimum</i>	<i>Mean</i>	<i>Maximum</i>	<i>Standard Deviation</i>
Central	495	1,309,231	8,626,193	2,131,951
Emergent	273	207,090	998,544	284,613
Peripherals	18	11,644	390,874	40,034
<i>Number of documents</i>				
Central	681,804	2,617,595	8,626,193	2,409,770
Emergent	135,843	413,748	998,544	276,394
Peripherals	101	23,204	390,874	54,282
<i>Citações</i>				
Central	14,278,721	42,496,602	177,434,935	51,439,691
Emergent	1,660,840	5,335,987	13,772,961	3,453,155
Peripherals	1,082	183,303	2,938,841	442,128
<i>Citations per document</i>				
Central	7.44	18.39	24.56	5.25
Emergent	6.50	16.73	26.10	5.17
Peripherals	2.68	12.67	35.55	5.05
<i>HDI</i>				
Central	0.723	0.883	0.922	0.062
Emergent	0.609	0.863	0.944	0.078
Peripherals	0.350	0.681	0.946	0.150

Among the analyzed indicators, there was a significant difference among the groups for h-index, number of documents and citations. Similarity was found between Central and Emergent groups for citations per documents indicator. In addition to these relation, we found that HDI of Central and Emergent groups also did not differ significantly, considering a 5%

significance level. This research is justified for contributing for methodological frameworks of Metric Studies, as we rank the countries through bibliometric indicators, performing a taxonomy.

Keywords: Metric Studies of Information; Bibliometric indicators; Taxonomy; Central, emergent and peripheral countries.

REFERENCES:

Guimarães, J. A. C. (2000). Perspectivas de ensino e pesquisa em organização do conhecimento em cursos de biblioteconomia do Mercosul: uma reflexão. In: Encuentro de EDIBCIC, 5., Granada. *La formación de profesionales e investigadores de la información para la sociedad del conocimiento: Actas*. Granada: Universidad de Granada, 206-216.

Hair, J. F. et al. (2009). *Multivariate Data Analysis*. 7 ed. Upper Saddle River: Prentice.

Mugnaini, R., Jannuzzi, P. & Quoniam, L. (2004). Indicadores bibliométricos da produção científica brasileira: uma análise a partir da base Pascal. *Revista Ciência e Informação*. 33, 123-131.

Oliveira, E. F. T & Grácio, M. C. C. (2011). Indicadores bibliométricos em ciência da informação: análise dos pesquisadores mais produtivos no tema estudos métricos na base Scopus. *Perspectivas em Ciência da Informação*, 16(4), 16-28.

PNUD (2015). Ranking IDH Global 2014. Acesso em: 20/02/2016. Available at: <http://www.pnud.org.br/atlas/ranking/Ranking-IDH-Global-2014.aspx>.

PNUD (2016). Desenvolvimento Humano e IDH. Access: 15/03/ 2016. Available at: <http://www.pnud.org.br/IDH/DH.aspx>.

Thomaz, P. G, Assad, R. S. & Moreira, L. F. P. (2011). Uso do fator de impacto e do índice h para avaliar pesquisadores e publicações. *Arquivo Brasileiro de Cardiologia*, 90-93.

Vanz, S. & Stumpf, I. (2010). Procedimentos e ferramentas aplicados aos estudos bibliométricos. *Informação e Sociedade*, 20(2), 67-75.



Uncriticized citation process of the indicators like social contagion - a case study of the success rate of commercialization of the public R&D in South Korea

Jinseo Park^{*}, Sun-Woo Kim^{**}, June Young Lee^{**} and Tae Ho Song^{**}

* jayoujin@hanmail.net

Department of Scientometric Research, Korea Institute of Science and Technology Information, 66 Hoegiro, Dongdaemun-gu, Seoul, 02456 (South Korea)

** kimsw@stepi.re.kr; road2you@kisti.re.kr; thsong@pusan.ac.kr

Entrepreneurship Research Unit, Science and Technology Policy Institute, Sejong National Research Complex 370, Sicheong-daero, Sejong-si, 20147 (South Korea); Department of Scientometric Research, Korea Institute of Science and Technology Information, 66 Hoegiro, Dongdaemun-gu, Seoul, 02456 (South Korea); School of Business, Pusan National University, Busandaehak-ro 63beon-gil, Geumjeong-gu, Busan, 46241 (South Korea)

ABSTRACT

One of the key problems in Korea's Science and Technology (S&T) policy is that "the success rate of commercialization (SRC) of public R&D" has been too low. For this reason, the Korean government has adopted "the SRC of public R&D" as an important performance or goal indicator. In 2009, the National Science & Technology Commission, which is the highest decision-making body for S&T policies in Korea, pointed to the fact that Korea's SRC of 24.6%-41% was excessively low compared to the rates in the United Kingdom (70.7%), United States (69.3%), and Japan (54.1%) and claimed this to one of the current challenges facing public R&D. Thereafter, the enhancement of the SRC of public R&D has functioned as a major policy goal in Korea's S&T policy. In the mass media, the SRC of public R&D as an indicator and comparative figures contrasting the rate to the aforementioned indicator and related figures have been deployed as a stylized fact, as representing the inefficiency of Korea's public R&D.

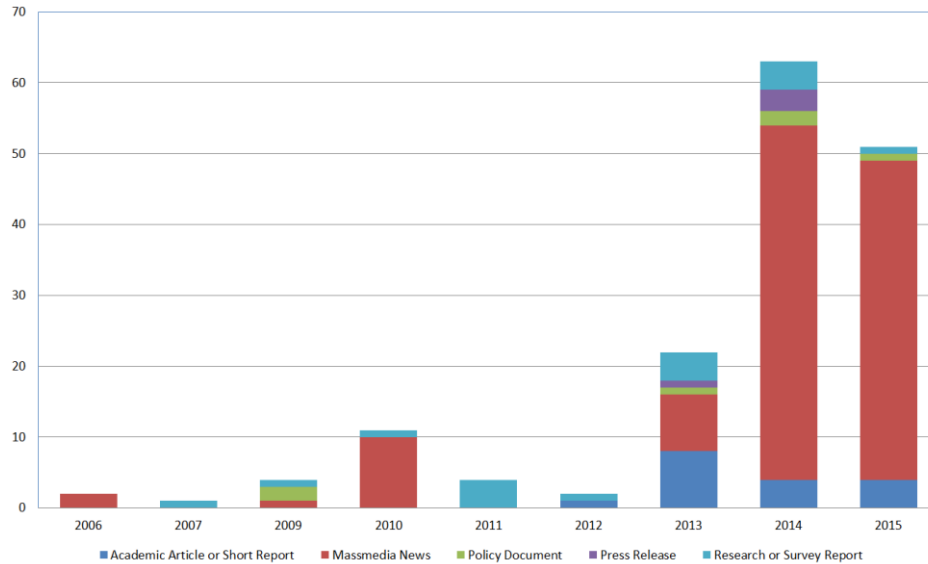
In this study, we examine the citation process of how the indicators and figures for the SRC of public R&D have been disseminated through Korea. The report of the National Science & Technology Commission did not include a clear definition of commercialization, the measurement method, or the references on which the figures for the United Kingdom, United States, and Japan were based. As observed by Latour (1987) and Gläser & Laudel (2007), a positive modality operates, whereby the peripheral conditions and the context in which the indicators and figures were generated are erased, and the vague indicators and figures acquire the status of a stylized fact.

DATA AND METHODS

We collected, through web search engines including Google, Naver, and Daum, all relevant documents that compared the figures for the SRC of public R&D in the United Kingdom, the United States, and Japan with the figures for Korea. As a result of our searches, we found 161 documents (Figure 1). We classify citation modes into three types. Type 1 are cases in which the reference is clearly provided in the sentence or paragraph that has the cited figures. Type 2 are cases in which the figures are cited using quotation marks, but the reference is absent and only the source quoted is given by naming a specific institution or person (such as "according to Ministry of . . ." or "according to Dr . . ."). Type 3 are cases in which there are no quotation marks, no indications of citation, and no source quoted. Whereas Type 2 is a citation mode that

borrowing authority from a specific institution or individual to validate an argument, Type 3 is a citation mode in which the status of a stylized fact is conferred on the figures. We constructed the citation links for all documents and analyzed the citation patterns.

Figure 1. Number of documents that compared the figures for the success rate of the commercialization of public R&D



RESULTS

Figure 2 presents the citation mode and annual trends for the 117 newspaper items. There is only a single newspaper item that contains an accurate reference for the given figures (Type 1). As for the majority, there were 58 newspaper items that cited a specific institution or individual without a reference (Type 2), and there were 58 items that lacked any marks of citation or source quoted.

Figure 2. The citation mode and annual trends for newspaper items

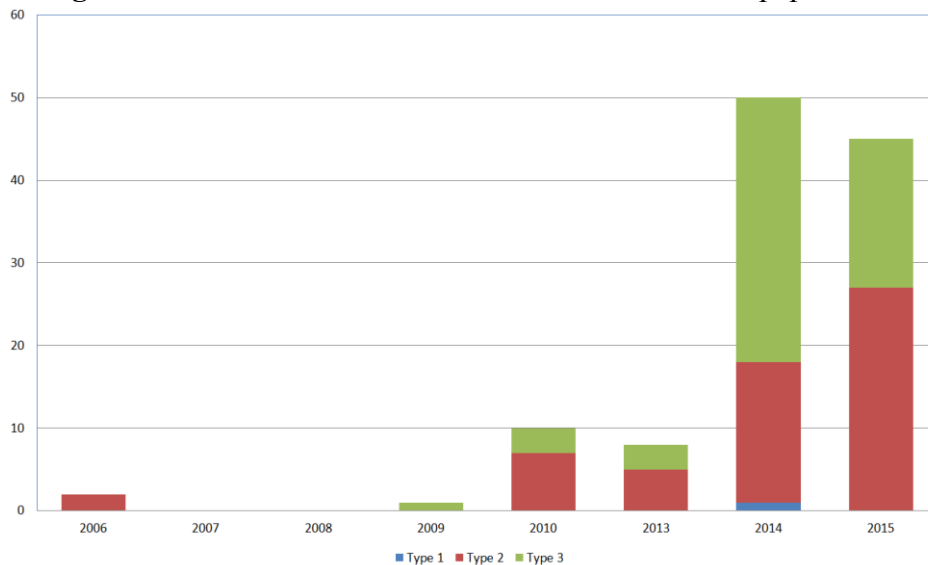


Figure 3. The citation process in the total of 161 cases

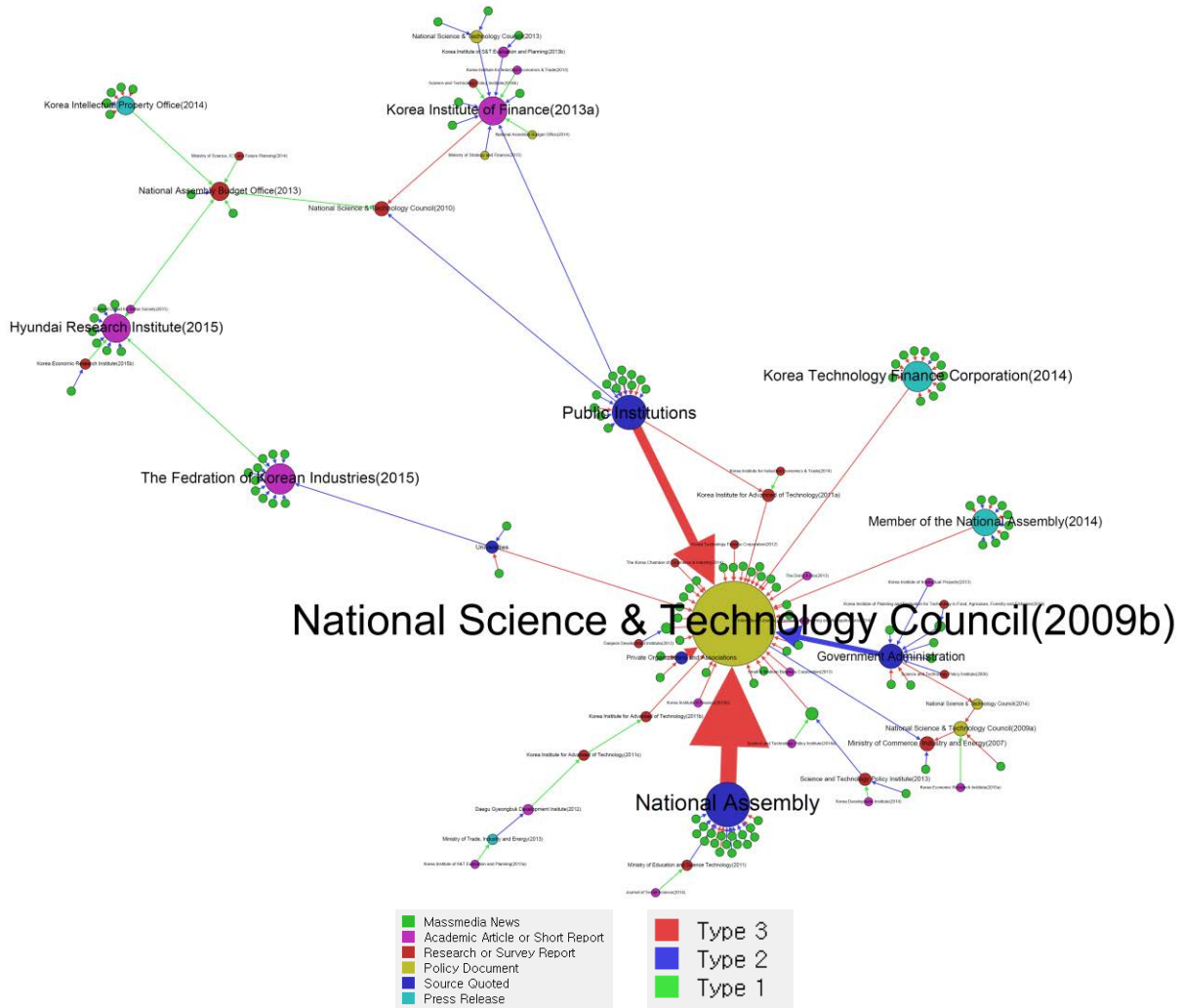


Figure 3 is a schematic diagram of the citation process for the 161 cases we examined. The nodes and the links are color-coded. If the source quoted in a newspaper item is an institution or individual, or if the item is a column that was written by an expert instead of a journalist, we classify the institution to which the source quoted or the column contributor belongs into the categories of public institutions, government administration, National Assembly (or a member of the National Assembly), private organizations and associations, and universities, and a link is made to the reference that we deduce to be the original citation. If it is a case in which the reference cannot be deduced, we regard this as a link to the 2009 report by the National Science & Technology Commission, which was the first to present the comparative figures for this indicator.

In conclusions, the overall process by which vague indicators and figures were consecutively cited is similar to cascade networks or the spread of an epidemic disease (Easley & Kleinberg, 2010). Furthermore, we have found that press releases made by the government and public institutions and reports written by experts who belong to the government and public institutions were important carriers through which vague indicators and figures came to be disseminated into mass media.

REFERENCES

Easley, D. & Kleinberg, J. (2010). *Networks, Crowds and Markets - Reasoning about a Highly Connected World*. Cambridge University Press.

Gläser, J. & Laudel, G. (2007). Chapter 5. The Social Construction of Bibliometrics Evaluations. In R. Whitley, & J. Gläser (Eds.). *The Changing Governance of the Sciences - The Advent of Research Evaluation Systems* (pp. 101-123). Springer.

Latour, B. (1987). *Science in Action - How to follow scientists and engineers through society*. Harvard University Press.



A comparative analysis of Western Europe and Latin America based on social and scientific indicators¹

Renata Cristina Gutierrez Castanha* and Maria Cláudia Cabrini Grácio**

*regutierrez@gmail.com

**cabrini@marilia.unesp.br

Information Science Department, UNESP – Univ Estadual Paulista, Hygino Muzzi Filho, 737, Marília – SP, 17525-000 (Brazil).

ABSTRACT

Starting from the idea of central and peripheral areas (Shils, 1961), that a society is constructed by areas that relate to each other unequally and that it's possible to identify the distance and direction of the interactions between them, this paper aims to collaborate to understand the ties inside and out the peripheral and the central areas, to know: Latin America (LA) and Western Europe (WE), respectively, through comparison of their social and scientific indicators, in order to visualize their similarities and dissimilarities, also to contribute to studies of dynamics of science in these regions. These regions were chosen in order to compare their behavior, being different in its classification, once LA is a peripheral area and WE is a central one. From World Bank Indicators and Scimago JR, it was recovered the indicators: GDP per capita (US\$), Researchers in R&D (per million people), Research and development expenditure - RDE (% of GDP), Human Development Index - HDI, Citable Documents, Cites per Document, Cited Documents and % of International Collaboration from 2000 to 2015. The data were classified as social and scientific indicators and their average values were organized in a spreadsheet. It was used Mann-Whitney test to compare the indicators from the areas and Pearson's correlation to analyse the association among them. Regarding the analyzed indicators, it was constructed both boxplot graphics and cluster analysis for the countries grouping, in order to verify, if the concept of centrality and geographic and social periphery of these two areas also reflects it in the scientific environment. Except for % of International Collaboration, both social and scientific indicators show much lower trends in LA in relation to WE, despite their similar behavior. The % of International Collaboration showed no difference in behavior between regions and is not correlated with any of the other analyzed indicators, meaning that this indicator is a scientific characteristic of the countries, independent of their scientific size or social or scientific infrastructure therein. Although, it is observed that in WE, the notion of centrality is not homogeneous.

¹ This work was supported by CNPq – Nacional Council of Technological and Scientific Development.

INTRODUCTION

The organization of scientific activities has its structure divided between two areas: central and peripheral, similar to social structures. The central zone is innovative and generating knowledge, having better means of control and development, producing greater and more significant amount of knowledge, guiding and establishing the criteria by which scientific communities will be evaluated. Developing countries would be mostly in this situation, a fact compounded when their native language differs from that used by the center, as ideal model in view of governments, institutions and individuals in these peripheral countries is to match up the center for purposes recognition, and so use their efforts to overcome the distance separating them (Shils, 1961; Nye, 2011).

Starting from the idea of central and peripheral areas (Shils, 1961), that a society is constructed by areas that relate to each other unequally and that it's possible to identify the distance and direction of the interactions between them, this paper aims to collaborate to understand the ties inside and out the peripheral and the central areas, to know: Latin America (LA) and Western Europe (WE), respectively, through comparison of their social and scientific indicators, in order to visualize their similarities and dissimilarities, also to contribute to studies of dynamics of science in these regions. These regions were chosen in order to compare their behavior, being different in its classification, once LA is a peripheral area and WE is a central one.

METHODOLOGICAL PROCEDURES

It was chosen the countries where, together, sum more than 90% of all scientific production, resulting in Brazil, Mexico, Argentina, Chile and Colombia for LA and United Kingdom, Germany, France, Italy, Spain, Netherlands, Switzerland, Sweden, Belgium and Austria for WE.

From World Bank Indicators² and Scimago JR³, it was recovered the indicators: GDP per capita (US\$), Researchers in R&D (per million people), Research and development expenditure - RDE (% of GDP), Human Development Index - HDI, Citable Documents, Cites per Document, Cited Documents and % of International Collaboration from 2000 to 2015. The data were classified as social and scientific indicators and their average values were organized in a spreadsheet. It was used Mann-Whitney test to compare the indicators from the areas and Pearson's correlation to analyse the association among them. Regarding the analyzed indicators, it was constructed both boxplot graphics and cluster analysis for the countries grouping, in order to verify, if the concept of centrality and geographic and social periphery of these two areas also reflects it in the scientific environment.

PRESENTATION AND DATA ANALYSIS

Table 1 and boxplots presented in Figure 1 show the behavior of analyzed indicators in a comparative way between LA and WE.

² <http://www.worldbank.org>

³ <http://www.scimagojr.com>

Table 1. Average and standard deviation for indicators, comparing regions LA and WE.

Indicator	LA		WE	
	Average	St. deviation	Average	St. deviation
GDP per capita (US\$)**	7826.96	1911.49	40465.69	9848.67
Researchers in R&D (per million people) **	458.84	292.34	3472.57	1021.14
RDE (% of GDP) **	0.49	0.34	2.10	0.71
HDI**	0.76	0.05	0.89	0.02
Citable Documents**	12810.69	13037.60	55472.05	38732.79
Cites per Document**	11.41	1.15	18.40	2.83
Cited Documents**	8798.97	9004.11	43251.79	30116.66
% of International Collaboration	42.38	10.50	44.78	6.70

** significance 1%

Except for % of International Collaboration indicator, all remaining indicators analyzed showed significant differences between LA and WE regions, with the social and scientific indicators of the LA area on average shaping up far below the WE area.

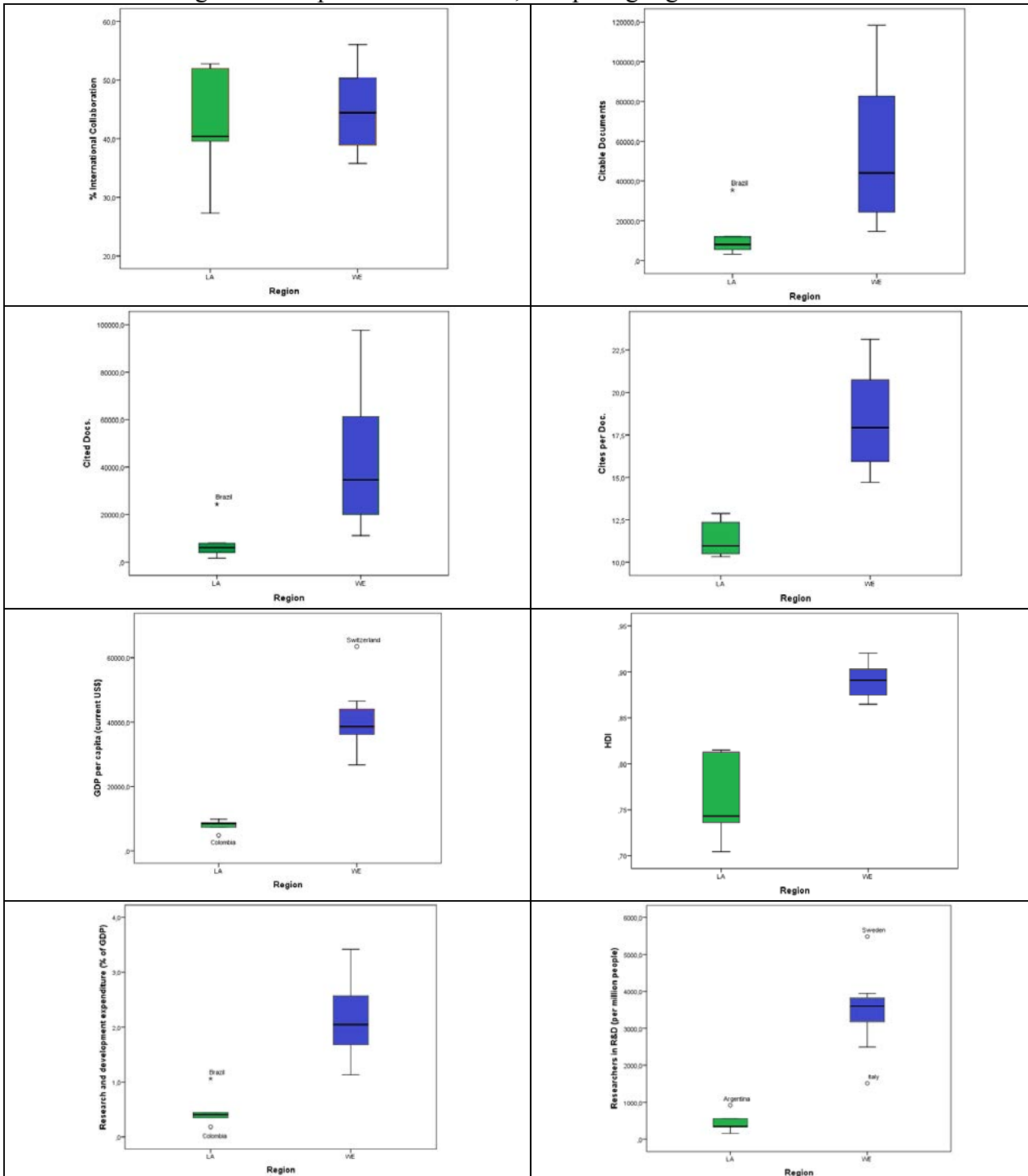
HDI of LA presents a shapeless variation between countries, while WE is more uniform and higher. In relation to Researchers in R&D, AL has far fewer compared to the WE and Argentina stands out as an outlier in LA.

Although, despite the larger number of Researchers in R&D to be present in WE, it occurs in many different ways. Italy has a total of researchers far below the average of countries in the area. The GDP per capita (current US\$) of LA is almost three times lower than the WE countries. It highlights Colombia and Sweden, upper and lower outliers, respectively.

The amount of both Cited and Citable Documents shows similar behavior, with a very low total in LA, highlighting Brazil as upper outlier, while in WE the boxplot covers a very wide range, probably due to the total of investigated countries. WE has higher number of Citations per Document than LA.

From the analysis of Pearson's correlations for indicators of LA area, there was a strong positive correlation between RDE and Citable ($r=.982$) and Cited Documents of the area ($r=.989$).

Figure 1. Boxplots for indicators, comparing regions LA and WE.



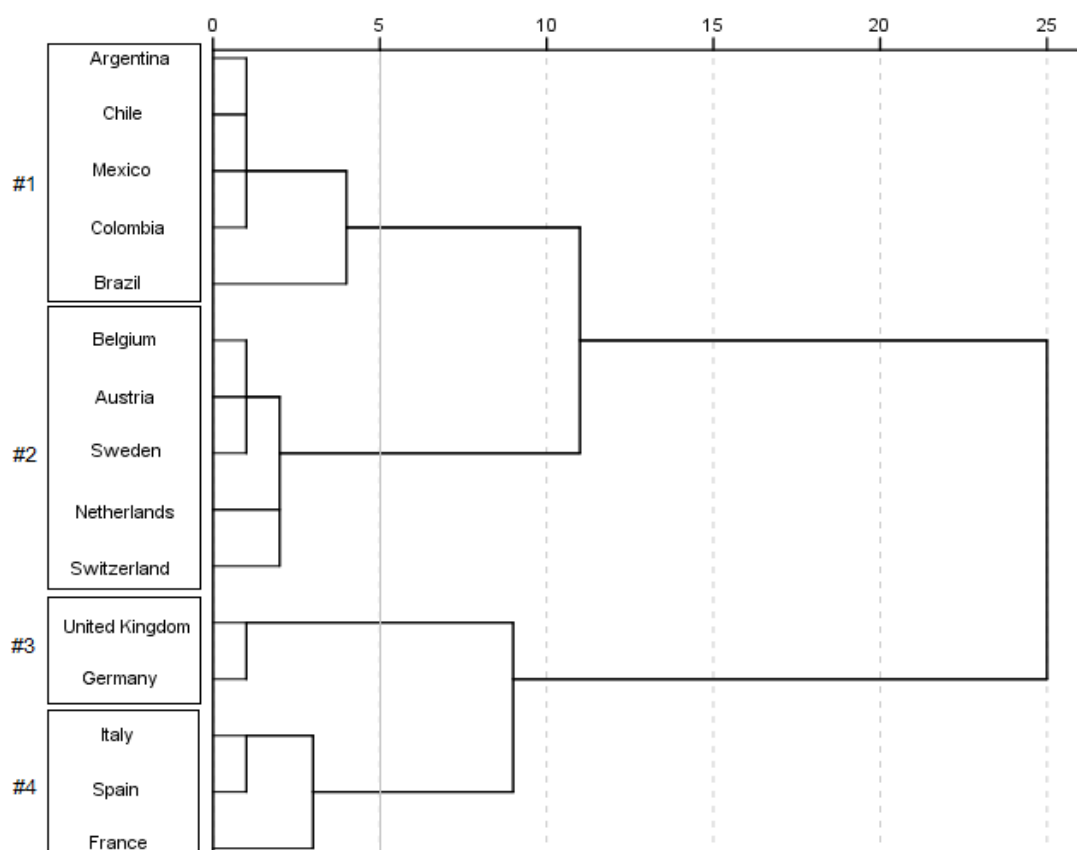
Regarding WE, also the higher is its GDP per capita (US\$), the larger is the number of citations received by documents and also their RDE. It is observed that both LA and WE have similar behavior regarding the analyzed indicators.

Whereas there was no significant difference between % of International Collaboration of the regions, Pearson's correlations between this indicator and the others were calculated

regardless of the area. The figures showed that the % of collaboration is not correlated to any of the other analyzed indicators.

Figure 2 shows the cluster, where it is observed that the countries of LA were grouped into a single group (# 1), while WE were divided into three different groups (# 2, # 3 and # 4), showing that even within the WE the notion of centrality is not homogeneous.

Figure 2. Cluster of social and scientific indicators.



FINAL CONSIDERATIONS

Except for % of International Collaboration, both social and scientific indicators show much lower trends in LA in relation to WE, despite their similar behavior. The % of International Collaboration showed no difference in behavior between regions and is not correlated with any of the other analyzed indicators, meaning that this indicator is a scientific characteristic of the countries, independent of their scientific size or social or scientific infrastructure therein. Although, it is observed that in WE, the notion of centrality is not homogeneous.

REFERENCES

- Nye, M. J. (2011). *Michael Polanyi and his generation: Origins of the social construction of science*. University of Chicago Press.
- Shils, E. (1961). Centre and periphery. In *The logic of personal knowledge: Essays presented to Michael Polanyi* (pp. 117-131). London: Routledge & Kegan Paul.



Indicators of endogamy and reciprocity in phd theses assessments¹

Lourdes Castelló-Cogollos^{*,***}; Rafael Aleixandre-Benavent^{*,**},

Rafael Castelló-Cogollos^{***}

* lourdes.castello@uv.es

UISYS-Universitat de València. Plaça Cisneros, 4. 46003-València, Spain

** rafael.aleixandre@uv.es

INGENIO (CSIC-Universitat Politècnica de València).

*** rafael.castello@uv.es

Departament de Sociologia i Antropologia Social. Facultat de Ciències Socials. Universitat de València Av. Tarongers, 4b. 46021 València, Spain

INTRODUCTION

In Spain, in the evaluation of the theses several academics contribute: supervisors, chairs and memberships of the academic assessment boards (AAB). Supervisors propose members of AAB and the PhD commissions from universities select these academics as members of the definitive AAB. This fact generates evident relationships between members that can be analysed and measured using indicators. Our objective is to propose indicators describing relations between supervisors involved in the academic assessment in the Spanish theses. These indicators, when applied to a complete set of theses in a specific field, can provide insights on relations of patronage and academic power between academics and contribute to draw a picture of the field structure from a sociological point of view.

METHOD

We use data from 3,413 theses obtained from TESEO, database of the Ministry of Education, Culture and Sports, which includes the Spanish theses defended and approved after its evaluation. The search was limited to the theses defended in the universities with degree in Sociology, indexed by UNESCO code related to Sociology.

We propose two indicators to measure the level of relation between supervisors: a **Endogamy Index (EI)**, which measures the level of closing of each supervisor in relation to the participation of different members to the AAB in supervised theses; and a **Reciprocity Index (RI)**, which measures the level of symmetry in the reciprocity between each two academicians in the participation to the AAB of the respective theses.

A) To obtain the EI, we propose to combine two sub-indexes.

¹ Supported by MINECO (CSO2012-39632-C02-01) and Genalitat Valenciana (GVPrometeo2013-041).

1. Closedness Index (CI), which measures the level in which a supervisor closes the participation in the AAB to a reduced number of academicians. We supposed that each AAB is compound for five members.

a. First, we establish a maximum number of academic participants in all the AAB of the theses of a same supervisor (MaxM): a maximum of two members repeated in all the AAB and three members that are always different. We obtain this maximum adding two to the number of theses (Nth) multiply by three. Therefore, $MaxM=2+(3*Nth)$.

b. Second, we calculate a ratio between the number of members participants (NM) and the maximum of members (MaxM). When $NM=MaxM$, the ratio will be 1 indicating the maximum openness. Therefore, if we discount this ratio to 1, we obtain a no-normalised² closedness index (CIa). So, $CIa=1-(NM/MaxM)$.

c. Third, we calculate the maximum of no-normalised closedness (MaxCIa), which we obtain with the division of 5 (minimum of participants in all the AAB) for the maximum number of possible members (MaxM). Therefore, $MaxCIa=1-(5/MaxM)$.

d. Finally, we obtain the CI as the ratio between CIa and MaxCIa. So, $CI = CIa/MaxCIa$.

2. Concentration Index (CoI), which measures the level in which a supervisor concentrates the majority of participations in some few academicians.

a. First, we calculate the geometrical average of the relative participations of each member of the AAB over all the participations in the AAB of the theses of the supervisor –Relative Participation (RP)–.

b. Second, we calculate the value of the proportion that supposes a theoretical equal participation of all the participants in the total of participations –Equitable Participation (EP)–, which is the ratio between 1 and the number of members (NM).

c. Finally, to obtain the CoI, we calculate the ratio between EP and RP. When $RP=EP$, we have an equitable distribution. Therefore, if we subtract 1 to the ratio obtained, we have an indicator of concentration. So, $CoI=(PE/PR)-1$.

3. Endogamic Index (EI) is the result to combine the CI with the CoI, so: $EI=C1*(1+CoI)$.

B) To obtain the RI, which measures the level in which two supervisors keep a symmetrical relation in his mutual participations in the AAB of the other, we propose the following steps:

² Since the maximum depends of the number of theses supervised.

1. Participation Ratio (PaR): number of participations over the total theses. It's 1 when a supervisor has participated in all the theses supervised by the other.
2. Reciprocity of Participation Index (RPI): it's the ratio of the minimum PaR of the two supervisors divided by their maximum. It's 1 when the presence of a supervisor in the AAB of the theses of the other is the same for the two supervisors, and therefore it indicates maximum symmetry.
3. Presidencies Ratio (PrR): number of presidencies of one supervisor in the AAB of the other over the total of participations in the AAB of the other.
4. Reciprocity of Presidencies Index (RPrI): it's the minimum of the PrR of the two supervisors divided by their maximum. This index is 1 when the presences of the two supervisors as presidents of the AAB of the other are identical.
5. Reciprocity Index (RI): it's the arithmetical average of RPI and RPrI.

RESULTS

The **Table 1** shows an example of the EI where we can see that the supervisor 2 (S2) is the highest EI followed by the S3. However, the S3 has highest CoI. The S4 has lowest EI. So, S2 and S3 are the most inbred supervisors; that is, often they invite the same people to join the AAB of their theses.

Table 1. Process to obtain the EI.

Supervisors	Nth	NM	MaxM	CIA	MaxCIa	CI	TR	RP	EP	CoI	EI
S1	14	27	44	38.6%	88.6%	43.6%	51	3.1%	3.7%	20.4%	52.5%
S2	12	17	38	55.3%	86.8%	63.6%	45	4.4%	5.9%	34.3%	85.5%
S3	21	31	65	52.3%	92.3%	56.7%	87	2.4%	3.2%	37.1%	77.7%
S4	24	67	74	9.5%	93.2%	10.1%	102	1.3%	1.5%	18.6%	12.0%

The **Table 2** shows an example of the RI where we can see that the academic pair E has the greatest reciprocity followed by the pair C. In the RPrI, the academic pairs B and D obtain a zero level, to indicate that their relations isn't symmetric; therefore, their relationship are of domination, power of an academic over another.

Table 2. Process to obtain the RI.

Academic pairs	Participation number	Chairs	Total theses supervised	PaR	RPI	PrR	RPrI	RI
A1 → A2	11	9	14	78.6%	78.6%	81.8%	73.3%	76.0%
A2 → A1	5	3	5	100.0%		60.0%		
B1 → B2	10	10	12	83.3%	48.0%	100.0%	0.0%	24.0%
B2 → B1	6	0	15	40.0%		0.0%		
C1 → C2	9	8	19	47.4%	82.1%	88.9%	80.4%	81.2%
C2 → C1	7	5	18	38.9%		71.4%		
D1 → D2	8	0	19	42.1%	67.4%	0.0%	0.0%	33.7%
D2 → D1	5	5	8	62.5%		100.0%		
E1 → E2	7	6	9	77.8%	71.4%	85.7%	93.3%	82.4%
E2 → E1	5	4	9	55.6%		80.0%		

CONCLUSIONS

These indexes can provide an insight of the groups configuration by the theses supervisors in relation to the formation of the AAB.



SCIENTIFIC IMPACT INDICATORS: A COMPARATIVE STUDY OF BRAZILIAN JOURNALS' IMPACT FACTORS

Cátia Cândida de Almeida* and Maria Cláudia Cabrini Grácio**
caticandida@gmail.com

Doctoral student in Information Science at UNESP – Univ Estadual Paulista and engineering professor, UniSALESIANO – Centro Univ Católico Salesiano Auxilium, 3821, Rod. Teotônio Vilela, Araçatuba, 16016-500(Brazil)

** cabrini@marilia.unesp.br
UNESP–UnivEstadual Paulista, 737 Hygino Muzzi Filho Avenue, Marília, 17525-900(Brazil)

ABSTRACT

In this context science social, this study aims to compare the IF of Brazilian journals, simultaneously present in Scopus and SciELO databases, in order to contribute to the discussion on the need to contextualize the use of IF as a quality indicator of scientific journals for the science denominated peripheral. The choice of multidisciplinary database Scopus is justified as this basis presents the highest number of indexed Brazilian journals.

INTRODUCTION

Scientific assessment is an important activity to understand the scientific behavior of different areas and countries, as recognized by the scientific community.

Among the bibliometric indicators that subsidize the assessment of scientific production, the impact factor (IF) created by Eugene Garfield in 1955, stand out for assessing the impact of scientific journals. IF can be understood as the average number of citations a scientific journals has in a given period.

IF is a widely used indicator as it is simple to be calculated and understood, however, its use has received considerable critics, because of its indiscriminate use due to its easy to understanding (Glänzel & Moed, 2002). Another limitation refers to the predominance of English language journals (Fernandes-Llimós, 2003).

Impact assessment of journals has been usually carried out from large databases recognized worldwide, which index mainstream science. Mainstream science is considered the ones indexed in major databases, whereas peripheral science, the less visible in international science (Velho, 1985). Despite the great importance of these bases, we also consider necessary to analyze the national and regional databases, in order to trace the real situation of scientific production in countries considered peripheral, such as Brazil.

In this sense, we highlight the Brazilian database named SciELO, created in 1996, which constitutes a network of journal collections published on the Internet, with open access, in order to promote visibility and access of scientific publications from Latin America and the Caribbean (Meneghini, 2003).

In this context, this study aims to compare the IF of Brazilian journals, simultaneously present in Scopus and SciELO databases, in order to contribute to the discussion on the need to contextualize the use of IF as a quality indicator of scientific journals for the science denominated peripheral.

The choice of multidisciplinary database Scopus is justified as this basis presents the highest number of indexed Brazilian journals.

METHODOLOGY

Initially, based on the study of Gracio and Oliveira (2014), we selected two scientific areas in Brazil: Dentistry, with high recognition in mainstream science; and Social Sciences, with low relative recognition in mainstream science. Through consulting the portal SCImago Journal & Country Rankings and SciELO database, we retrieved the journals from these two scientific areas, simultaneously present in both bases. From this criterion, we retrieved five Brazilian journals in Dentistry and five journals in Social Sciences. For each of the journals, we calculated the IF from both databases indexed in the years 2011, 2012 and 2013. Thus, using Excel software, IF for Scopus and SciELO were calculated with reference to 2014.

RESULTS AND DISCUSSION

Table 1 shows the IF calculated for Dentistry, for Scopus and SciELO.

Table 1. IF comparison for Dentistry area

Journals	Scopus IF	SciELO IF
Brazilian Dental Journal	1.12	0.34
Brazilian Oral Research	1.09	0.35
Journal of Applied Oral Science	1.02	0.26
Brazilian Journal of Oral Sciences	0.20	0.39
Dental Press Journal of Orthodontics	0.15	0.10

It is observed that in Dentistry, except for the Journal of Applied Oral Science, all journals have IFs in Scopus well above the IF observed for SciELO database. This behavior shows that this area has greater impact on mainstream science than in local science, suggesting high international recognition of the Brazilian science in the area.

Table 2 presents the calculated IF for the area of Social Sciences, for Scopus and SciELO.

Table 2. IF comparison for Social Sciences area

Journals	Scopus IF	SciELO IF
Saúde e Sociedade	0.34	0.53
Ambiente & Sociedade	0.30	0.15
Revista Brasileira de Ciências Sociais	0.17	0.32
Sociologias	0.17	0.24
Perspectivas em Ciência da Informação	0.13	0.12

In the area of Social Sciences, it is observed that except for the journal Ambiente & Sociedade, all others have their IF at SciELO higher than the IF at Scopus. This behavior may be due to the fact that researchers in this field prefer to publish in national journals and books (MENEHINI, 2003), but also because the social area is context dependent, often addressing matters of more regional interest than worldwide interests.

CONCLUSION

We concluded that the assessment of scientific impact of the journals should not be based on one single bibliometric indicator, since there is a risk of obtaining a partial visualization of its scientific behavior. To know and understand the real behavior of scientific journals of countries considered peripheral, such as Brazil, one should also consider the national and regional databases, especially in relation to scientific areas that are context-dependent.

REFERENCES

- Fernández-Llimós, F. (2003). *Análisis de la cobertura del concepto de Pharmaceutical Care en fuentes primarias y secundarias de información*. Granada: Universidad de Granada.
- Glänzel W. & Moed F. H. (2002). Journal impact measures in bibliometric research. *Scientometrics*, 53, n.2, 171-193.
- Grácio, M. C. C.& Oliveira, E. F. T. (2014). Estudo comparativo teórico metodológico dos índices normalizados de citação: uma aplicação na ciência brasileira (1996-2007). *Em Questão*, 20(3), 9-26.
- Meneghini, R. (2003). O projeto Scielo (Scientific Electronic Library on Line) e a visibilidade da literatura científica "Periférica". *Química Nova*, 26, n.2, 155-156.
- Velho, L.M.L.S. (1985). Como medir a Ciência? *Revista brasileira de tecnologia*, 16, p.35-41.



Sub-fields of Library and Information Science in Turkey: A Visualization Study¹

Zehra Taşkın*, Güleda Doğan*, and Umut Al*

*{ztaskin; gduzyol; umutal}@hacettepe.edu.tr

Hacettepe University, Department of Information Management, 06800, Çankaya, Ankara, Turkey

ABSTRACT

This study aims to determine the LIS sub-fields studied in Turkey and compare them with world LIS literature through keywords and abstracts of 460 peer-reviewed articles published in two LIS journals in Turkey (TL and IW) using VosViewer visualization software. Results of the study reveals that although sub-fields of LIS studied in Turkey similar with World LIS literature, it also has its own characteristics. While bibliometrics and library science are common sub-fields; LIS field in Turkey differs by the means of sub-fields named as records management and reading habits.

INTRODUCTION

Library and information science (LIS) is an interdisciplinary field that combines the various subjects such as information technology, libraries, management and education. It has been experiencing significant transformation on its main subjects on the occasion of recent developments in digital technology. The main problem for this huge transformation is creating gap between developed and developing countries. Researchers who work in the developing countries have information retrieval problems according to their countries' geographical location, political isolations, limited funds or embargos (Johnson, 2007, p. 65; Tella & Issa, 2011, p. xx; Uzun, 2002, p. 22-23). These kinds of restrictions effect information behaviour of researchers. Under these circumstances, it becomes vital to reveal countries' information environments for each discipline. To reveal them, bibliometrics and social network analyses are the commonly used methods in the literature.

The main aim of this study is to determine the LIS sub-fields studied in Turkey and compare them with world LIS literature. To achieve this aim, two research questions below are addressed:

- What are the main subject clusters for Turkish LIS literature and how they connect to each other?
- Do Turkish LIS literature converge to international LIS literature in the meaning of subject clusters identified?

METHOD

Two refereed LIS journals of Turkey -*Turkish Librarianship (TL)* and *Information World (IW)*- are selected to analyze the field. *TL* is a quarterly journal of Turkish Librarians'

¹ This work was supported by in part by a research grant of the Turkish Scientific and Technological Research Center (115K440).

Association published since 1952 (*Turkish Librarianship*, 2016). *IW* is published since 2000 and appears twice a year (*Information World*, 2012).

265 peer-reviewed articles published in *TL* between 1996-2015 (*TL* began to publish peer reviewed articles in 1996) and 195 peer-reviewed articles published in *IW* until 2015 are chosen for evaluation. All required information such as titles, abstracts, keywords, etc. gathered through a database designed. *VosViewer* that is tool visualize bibliometric networks (VosViewer, 2016) is used for visualizing sub-fields of LIS literature in Turkey.

RESULTS

To determine sub-fields of LIS in Turkey, keywords and abstracts of the 460 articles indicated in the methodology analyzed. Table 1 shows keywords occur in at least six articles, that comes from more than half of the articles (246 articles). According to the keywords of articles, studies on records management, libraries (university, public or school), bibliometrics including citation analysis, reading habits have an important role in Turkish LIS literature. On the other hand, studies on scholarly communication including open access, user studies and information literacy have an effect in the field.

Table 1. Keywords that occur in at least 8 articles

Keyword	n
Turkey	26
records management	20
public libraries	19
university libraries	18
bibliometrics	17
libraries	16
information literacy	12
reading habits	11
citation analysis	10
knowledge management	10
open access	9
scholarly communication	8
user studies	8

Figure 1 contains 92 noun phrases having highest relevance score in 6147 noun phrases extracted from abstracts of 460 articles. Each of the four clusters represents a sub-field of LIS in Turkey. The largest cluster colored red at the upper left part of the map is well-separated from other three clusters refer to “bibliometrics/citation analysis” as prominent noun-phrases are “article”, “analysis”, “year”, “number”, “journal”, “publication period” and “database”.

Three clusters colored yellow, green and blue at the right part of the map are not well-separated from each other, especially two of them (blue and green clusters) at the right part of the map are connected to each other. Yellow cluster at lower middle part of the map is more likely to be the “records management” sub-field of LIS; with prominent noun phrases “system”, “organization”, “management”. Green and blue connected clusters at upper right part are considered as “library science” (green cluster) and “reading habit” (blue cluster) sub-fields of LIS. Prominent terms for “library science” cluster colored green are “service”,

REFERENCES

- Åström, F. (2002). Visualizing library and information Science concept spaces through keyword and citation based maps and clusters. In *CoLIS4 Proceedings of the Fourth International Conference on Conceptions of Library and Information Science* (Bruce, H., Fidel, R., Ingwersen, P. & Vakkari, P. Eds.) pp. 185-196. Colorado: Libraries Unlimited.
- Järvelin, K. & Vakkari, P. (1993). The evolution of library and information Science 1965-1985: A content analysis of journal articles. *Information Processing and Management*, 29(1), 129-144.
- Johnson, C.A. (2007). Library and information science education in developing countries. *International Information & Library Review*, 39(2), 66-71.
- Information World. (2012). Retrieved March 4th, 2016 from: <http://bd.org.tr/index.php/bd>
- Tella, A. & Issa, A.O. (2011). *Library and information science in developing countries: contemporary issues*. Hershey, PA: IGI Global.
- Turkish Librarianship. (2016). Retrieved March 4th, 2016 from: <http://www.tk.org.tr/index.php/TK>
- Uzun, A. (2002). Library and information science research in developing countries and eastern European countries: A brief bibliometric perspective. *International Information & Library Review*, 34, 21-33.
- Van Eck, N.J. & Waltman, L. (2011). Text mining and visualization using VOSviewer. *ISSI Newsletter*, 7(3), 50-54. Retrieved March 4th, 2016 from: <https://arxiv.org/ftp/arxiv/papers/1109/1109.2058.pdf>
- VosViewer. (2016). Retrieved March 4th, 2016 from: <http://www.vosviewer.com/>



Content words as measure of structure in the science space¹

Wout S. Lamers*

* w.s.lamers@cwts.leidenuniv.nl

Centre for Science and Technology Studies (CWTS), Leiden University, Leiden (The Netherlands)

ABSTRACT

We ask the question whether it is possible to construct maps of science and find disciplinary similarity structures based not on citation, but on the content of publications' titles and abstract. Using noun phrases and 33 disciplines made up of a total of 7 million publications from distinct groups of Web of Science subject categories, we generate discipline similarity structures and maps using cosine similarity of term occurrence vectors. We find that these structures are highly stable, robust against the removal of the 99.5% of lower-relevance noun phrases, and may be used to classify publications into disciplines with encouraging accuracy.

INTRODUCTION

Over the past few decades, several authors have argued that profound changes are taking place in the way new scientific knowledge is produced. Publications such as *The New Production of Knowledge* (Gibbons et al., 1994) and subsequent alternatives such as the triple helix model of industry, academia and government interaction (Leydesdorff & Meyer, 2006) have set off a debate on whether the disciplinary divisions in the science system are still an accurate representation of reality. Increasing collaboration between science and industry to address complex socio-scientific issues is prompting changes in the ways researchers and policymakers alike think about performance measures and evaluation of scientific output (Wagner et al., 2011). This increasing collaboration, both between and beyond scientific disciplines, increases the complexity of the science system. From a scientometrician's perspective, these developments pose significant challenges, and require an extension of our methodological toolbox.

The traditional tools used for gaining insight into the science system and its structure are various types of bibliometric mapping techniques, constructed most frequently using various types of citation data. For instance, global maps of science have been created at the journal level based on the Web of Science subject categories (Leydesdorff, Carley, & Rafols, 2013; Leydesdorff & Rafols, 2009) as well as at the publication level using direct citation relations (Waltman & van Eck, 2012). The maps produced by both methods allow researchers to decompose the body of science into disciplinary and sub-disciplinary structures. However, given that they rely on citation data for the positioning and clustering of nodes, they are limited in the sense that these disciplinary structures cannot subsequently be applied to classify, for instance, research proposals and grey literature, or any type of text that does not adhere to academic citation standards. Non-citation based methods, such as co-word mapping have been used to map limited journal sets (e.g. van den Besselaar & Heimeriks, 2006) and to describe research topics of citation clusters (e.g. Braam, Moed, & van Raan, 1991), but not to map the overall structure of science to the extent that citation-only methods have.

¹ The contents of this paper were previously published as a Master's Thesis at Utrecht University. I would like to thank Gaston Heimeriks (Utrecht University), Ingeborg Meijer (CWTS) and Ed Noyons (CWTS) for their supervision and guidance.

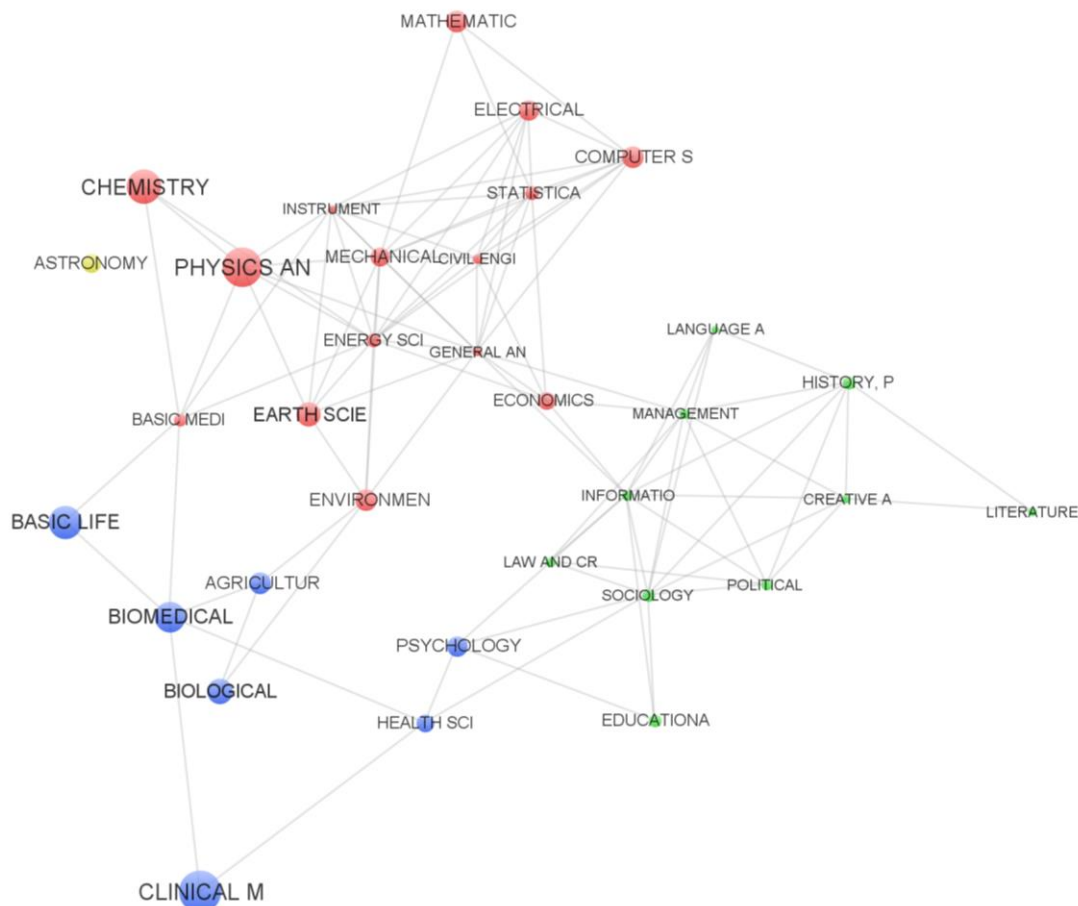
We ask the question whether it is possible to construct maps of science and find disciplinary similarity structures based not on citation, but on the content of publications' titles and abstract. If so, this may prove to be a valuable addition or expansion of traditional methods for generating maps of science and classifying publications.

METHOD

From the Web of Science database, we collected over seven million scientific articles spread over 33 groups of Web of Science subject categories representing disciplines. These groups of subject categories were developed by CWTS for the Science and Technology Indicators 2010 report of the Netherlands Observatory of Science and Technology (NOWT, 2010). The groups labelled “social sciences, interdisciplinary” and “multidisciplinary journals” were excluded, as were any publications with subject categories from multiple groups, in order to limit out data to monodisciplinary publications in the timeframe 2000-2010. From the enhanced Web of Science database available at CWTS, we also recovered the noun phrases used in these publications' titles and abstract, and for each publication we constructed a term occurrence vector listing whether noun phrases were present or absent in the text.

Discipline term occurrence vectors were computed by taking the mean of the discipline's associated publication's term occurrence vectors. Similarity scores between disciplines' term occurrence were subsequently computed using cosine similarity. We constructed a discipline similarity structure and use this structure to generate maps and clustering solutions. The results for the complete sample are presented in figure 1.

Figure 1. Map of 33 disciplines based on cosine similarities of term-occurrence within Web of Science subject category groups, 2000-2010. Size of nodes corresponds to the number of publications, the 100 highest-similarity pairs are displayed using the edges.



Comparing the results for all different years in our sample, we find a consistent clustering solution across the years representing three larger fields in science:

- A STEM cluster, with Basic medical sciences, Environmental sciences and technology, Economics, Mathematics and disciplines between those in figure 1.
- A life sciences cluster, with Basic life sciences, Psychology and Agriculture and food science at the borders, extending to Clinical medicine.
- A humanities cluster, left of Psychology and Economics, extending to Literature.
- Finally, Astronomy and astrophysics was consistently located in a cluster of its own.

Correlations of discipline pair scores compared over subsequent years were consistently high (>0.99). Comparing scores from more distant years showed a slight downward trend, but correlation remained high (>0.97).

We explored approaches for reducing the amount of noun phrases used in term occurrence vectors. By selecting the noun phrases whose occurrence frequency in disciplines differed most greatly from their occurrence in the complete sample, we reduced the amount of noun phrases used from 40 million to 20 thousand, while the resulting discipline pair similarity structures remained highly correlated (>0.999 comparing the complete sample with the high-relevant noun phrase sample).

Finally, we randomly split off a training sample and a test sample consisting of respectively 6.3 million publications and 0.7 million publications, recomputed the discipline term-occurrence vectors and similarity scores for the training sample, and classified the publications in the test sample according to the highest similarity score of their term occurrence vector to the discipline term occurrence vectors. This process was repeated twice for a total of 1.4 million classified publications. In 45% of the cases, publications were classified correctly, an encouraging result considering we are working with 33 disciplines. Classification performance dropped to 44% using only highly relevant noun phrases.

CONCLUSION

Our results indicate that we can indeed derive a stable structure of the science space from publications' title and abstract text. The classification shows that this structure can subsequently be put to use to place new publications into this structure with encouraging accuracy. This is an important conclusion, as so far methods for mapping the science space have been mostly restricted to citation data. These results open new avenues for research, potentially into the systematic assessment of novelty and new combinations in science.

REFERENCES

- Braam, R. R., Moed, H. F., & van Raan, A. F. J. (1991). Mapping of Science by Combined Co-Citation and Word Analysis. I. Structural Aspects. *Journal of the American Society for Information Science*, 42(1988), 252–266.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. SAGE Publications.
- Leydesdorff, L., Carley, S., & Rafols, I. (2013). Global maps of science based on the new Web-of-Science categories. *Scientometrics*, 94(2), 589–593.
- Leydesdorff, L., & Meyer, M. (2006). Triple Helix indicators of knowledge-based innovation systems. *Research Policy*, 35(10), 1441–1449.

- Leydesdorff, L., & Rafols, I. (2009). A Global Map of Science Based on the ISI Subject. *Journal of the American Society for Information Science and Technology*, 60(2), 348–362.
- NOWT. (2010). Science and Technology Indicators 2010. *Netherlands Observatory of Science and Technology*.
- van den Besselaar, P., & Heimeriks, G. (2006). Mapping research topics using word-reference co-occurrences: A method and an exploratory case study. *Scientometrics*, 68(3), 377–393.
- Wagner, C. S., Roessner, J. D., Bobb, K., Klein, J. T., Boyack, K. W., Keyton, J., ... Börner, K. (2011). Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature. *Journal of Informetrics*, 5(1), 14–26.
- Waltman, L., & van Eck, N. J. (2012). A new methodology for constructing a publication level classification system of science. *Journal of the American Society for Information Science and Technology*, 63(12), 2378–2392.



Study on the International and Domestic Subject Areas' Distributions¹

Wei Wenjie*, Yao Junlan ** and Liu Yuxian* **

* wjwei@lib.tongji.edu.cn

Tongji University Library, No. 1239 Siping Road, Shanghai (China),200092

** jlyao@lib.tongji.edu.cn

Tongji University Library, No. 1239 Siping Road, Shanghai (China),200092

*** yxliu@lib.tongji.edu.cn(Corresponding to Liu Y.X.)

Tongji University Library, No. 1239 Siping Road, Shanghai (China),200092

INTRODUCTION

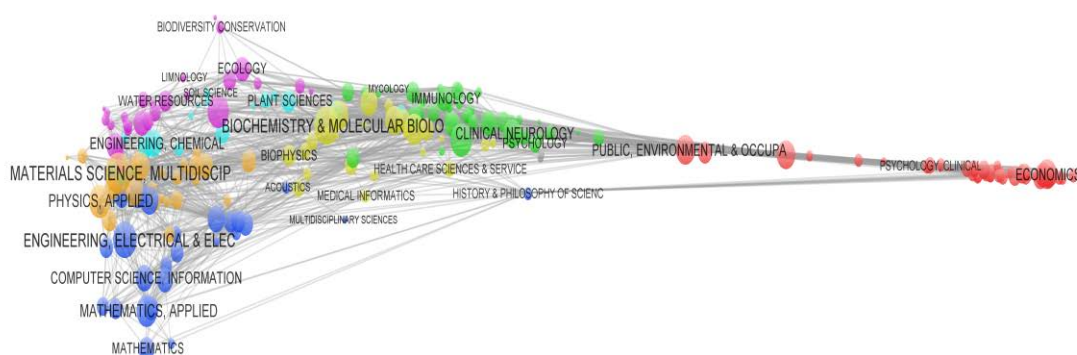
One journal may concern several topics. The journal hence can be divided into several subject areas. These subject areas are linked together via the journals. Based on this relation, we build two global maps: one uses the subject areas defined by the Web of Science, we call it as the international global map; the other use the subject areas defined by a Chinese database CNKI, we call it as the Chinese domestic global map. We overlap Tongji University's international publication data on the international global maps and Tongji University's domestic publication data on the Chinese domestic global map. We then look through the subject areas' distribution to see what is the difference of international research and domestic research of Tongji University.

INTERNATIONAL GLOBAL MAPS

Global map was invented by Leydesdorff & Rafols (2009). Based on citing and cited relation between the ISI subject areas, they made a factor analysis and decomposed scientific literature into 14 factors (later they considered social science and expended it to 19 factors). They visualized these factors and their citing or cited relations to get the global map. The subject areas are defined based on the journals' topics in the Web of Knowledge. Citation patterns was already considered when assigning subject areas to a journal. Moreover, the subject areas within a journal normally are more interrelated than that in different journals. Logically the global map created by co-occurrence of subject areas in one journal will represent the relation of subject areas more directly. So we create the co-occurrence metrics of subject areas and import into VoSviewer (Eck & Waltman 2009ab). Using VoSviewer's similarity algorithm, we cluster these subject areas and draw a network to show the international global map. The result network is showed in Figure 1. The distance between nodes reflects the similarity.

¹ This work was supported by This work was supported by NSFC via 71173154, National Social Sciences Foundation of China (08BZX076) and the Fundamental Research Funds for the Central Universities. Interdisciplinary Human and Social Science Foundation of Tongji University

Figure 1: International global map



Each node identifies a subject category of WOS. There are 226 categories totally in SCI-E and SSCI, extracted from 2014 JCR. We see nature sciences are on the left side, and social sciences are on the right side. The WOS categories are divided into 8 clusters by the clustering algorithm of VOSviewer. Though the global map drawn in Leydesdorff & Rafols (2009) has 19 clusters, both maps have similar structure.

Interestingly, two paths probable for blend of social sciences and nature sciences could be recognized. One path is category by category, which commences from fundamental categories such as mathematics and physics to engineering via material application, then voyages in sequence from ecology, biology, medicine, communal facilities to social sciences. Another is direct path from fundamental science to social science via history and philosophy of sciences.

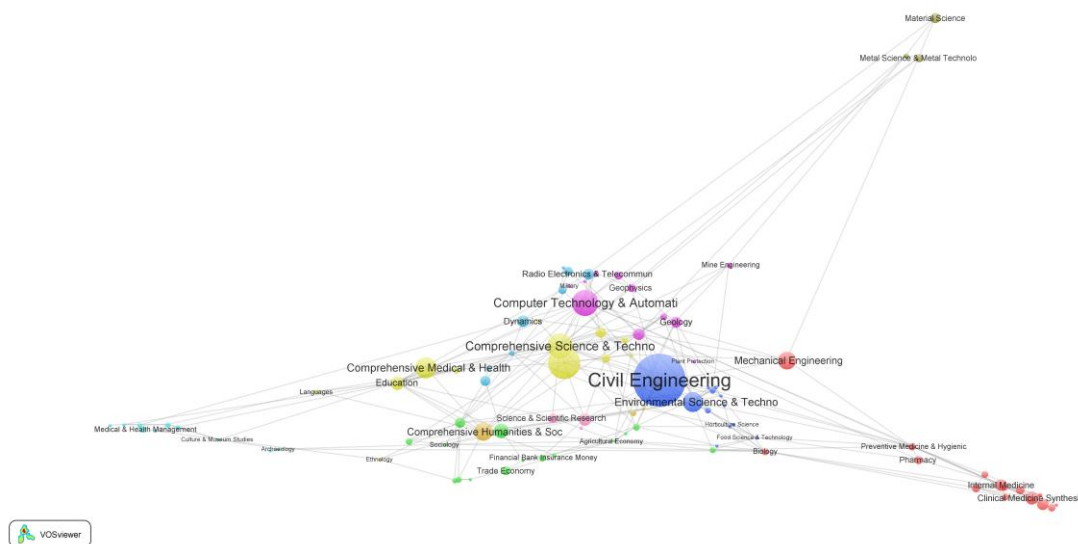
CHINESE DOMESTIC GLOBAL MAP

CNKI is the most comprehensive research platform available in China. The journals covered by the database also are assigned into different subject areas (the subject areas are different from that in WoS). Using the same method, we draw a Chinese domestic global map.

There are 18 clusters. 10 clusters are linked together, and 8 is isolated. The 10 linked clusters are showed in Figure 2. It is a triangle with social science at one angle, metal science at one angle and the medical science at one angle, in the middle social science are transferred to Engineering through Comprehensive Science and Technology and fundamental science such as mathematics, chemistry and physics.

indexed by CNKI, while material sciences, environment and ecology accounts for most in paper numbers indexed by WOS.

Figure 4: Simplified overlay map of Tongji CNKI categories from 2005 to 2014.



Science overlap map could visualize category distributions in different period, and enables contrast among institutions.

CONCLUSION

International global map and Chinese domestic global map has different characteristics, reflecting Chinese research is difference from the world. The difference may also be caused by Chinese special press policy and by a different catalogue system of the dataset we used. However, the researchers in same university with same priority in their subject areas have the similar interest, no matter their articles are published in international journals or domestic journals.

References

- Eck, N.J., & Waltman, L. (2009a). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523-538, doi:10.1007/s11192-009-0146-3.
- Eck, N.J.v., & Waltman, L. (2009b). How to normalize cooccurrence data? An analysis of some well-known similarity measures. *Journal of the American Society for Information Science and Technology*, 60(8), 1635-1651, doi:10.1002/asi.21075.
- Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology*, 60(2), 348-362, doi:10.1002/asi.20967.
- Rafols, I., Porter, A.L., & Leydesdorff, L. (2010). Science overlay maps: A new tool for research policy and library management. *Journal of the American Society for Information Science and Technology*, 61(9), 1871-1887, doi:10.1002/asi.21368.



Characteristics of Paper Publication by Major Countries Focusing on Journals

Naomi FUKUZAWA*

*fukuzawa@nistep.go.jp

Research Unit for Science and Technology Analysis and Indicators, National Institute of Science and Technology Policy (NISTEP), Kasumigaseki, Tokyo 100-0013 (JAPAN)

ABSTRACT

The number of journals and Open Access (OA) journals increased in the world. OA journals account for about 15% of all journals. The characteristics of paper publication were investigated for seven countries from the following five viewpoints: 1) whether the journal is OA, 2) whether the publishing country of the journal is the own country (domestic journal), 3) whether the paper is top 10% highly cited papers, 4) whether the paper is written in native language, and 5) whether the paper is international collaborated.

I categorized journals into four types; Domestic Non-OA, Foreign Non-OA, Domestic OA, Foreign OA. I analyzed the mean number of papers by the type of journal during the two periods of 2004-2006 and 2010-2012 and the ratio of those.

This report revealed that Japan had a unique trend where an increase in the number of foreign OA journals' papers contributed to an increase in the number of papers. In other countries, an increase in the number of foreign non-OA journals' papers contributed to an increase in the number of papers. A certain amount of Non-OA journals' paper was written in native language in the non-English-speaking countries. On the other hand, almost all OA papers written in English. The international collaboration ratio of OA journal's papers is high. In addition, OA journals' papers are better cited by foreign countries than Non-OA journals' papers.

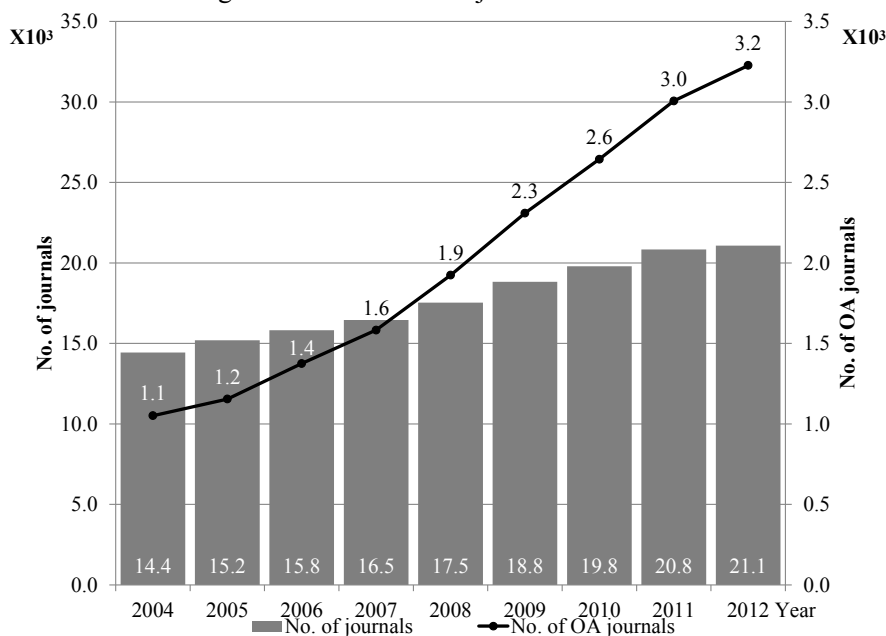
BACKGROUND AND PURPOSE

The number of papers is increasing globally. However, counting the number of papers alone does not clarify the countries' characteristics of paper publication. Recently, the number of journals and Open Access (OA) journals increased in the world (Figure 1). OA journals account for about 15% of all journals. When focusing on journals, it is not well studied how researchers publish papers. Therefore, this report aims at clarifying a part of the characteristics of paper publication in major countries by focusing on journals.

OA journals are journals where papers are published on the Internet, free of charge, and made accessible for anybody. When we publish a paper as OA, a method to make it free by the author(s) paying the article processing charges (APCs) (Gold OA), a method to list it on an institutional repository after a certain period of time after publication (Green OA), are mainly used (Wang, L. *et al.*, 2015). It was indicated that the amount of access by other researchers is higher in OA papers, resulting in an increase in the number of citations (Sothdeh *et al.*, 2015; Wang, *et al.*, 2015). It is possible that authors pay APCs by using research grants (Wang, L. *et al.*, 2015) to obtain such opportunities.

In this report, the characteristics of paper publication were investigated for seven countries from the following five viewpoints: 1) whether the journal is OA, 2) whether the publishing

country of the journal is the own country (domestic journal), 3) whether the paper is top 10% highly cited papers¹, 4) whether the paper is written in native language, and 5) whether the paper is international collaborated.

Figure 1: The number of journals in the world²

METHODS

The Elsevier Scopus Custom Data (February 19, 2015) was used as database. “Journal” was used as document type. The publication year was taken as the year. The number of papers was counted by the whole counting method. For counting OA journals, in the Journal Title List (2016 May) of Scopus, those listed as “DOAJ /ROAD Open Access³” at the journal level were counted as OA journals⁴. I categorized journals into four types; Domestic Non-OA, Foreign Non-OA, Domestic OA, Foreign OA.

RESULTS

First, I analyzed the mean number of papers by the type of journal during 2010-2012 and the ratio of those. Figure 2 and 3 show the results for all papers and for the top 10% highly cited papers, respectively. The number of OA journals’ papers was around 10% in each country.

¹ Papers within the top 10% share of the number of citations for each Scopus’s 27 field for each publication year.

² Data of all figures and tables in this report were collected conforming to the description in Methods.

³ DOAJ is a white list of open access journals (<https://doaj.org/about>). ROAD is service offered by the ISSN international Centre with the support of the Communication and Information Sector of UNESCO (<http://road.issn.org/en/contenu/purposes-road-project#.V3zL49KLRIY>).

⁴ The Journal Title List includes the OA journals in which all peer reviewed scholarly articles are online available without any restrictions and for which an APC has been paid, and the OA journals which do not charge an APC and are instead subsidized by other means.

Figure 2: Mean number of papers and their ratio for all papers

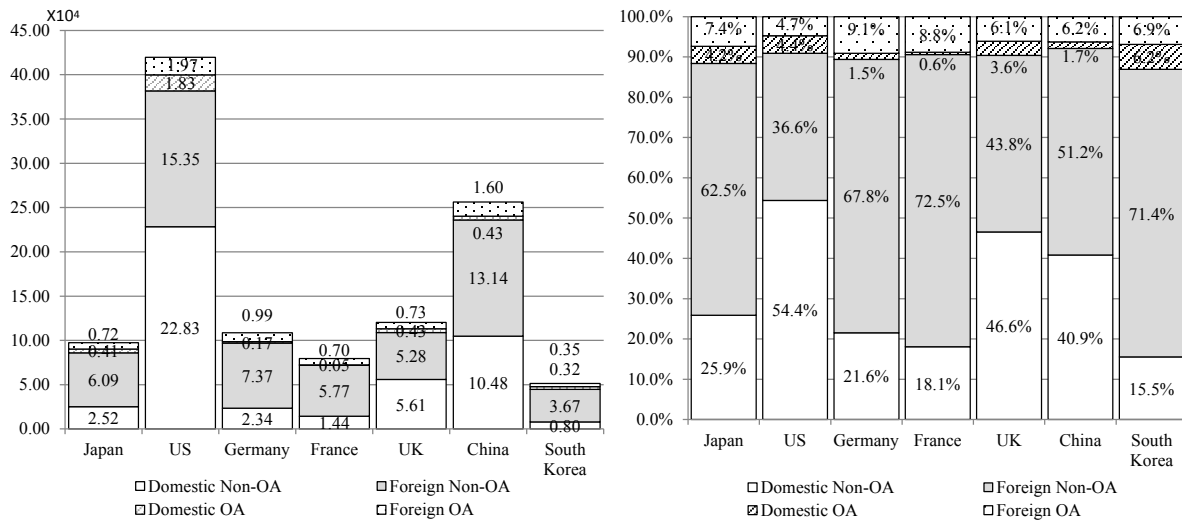
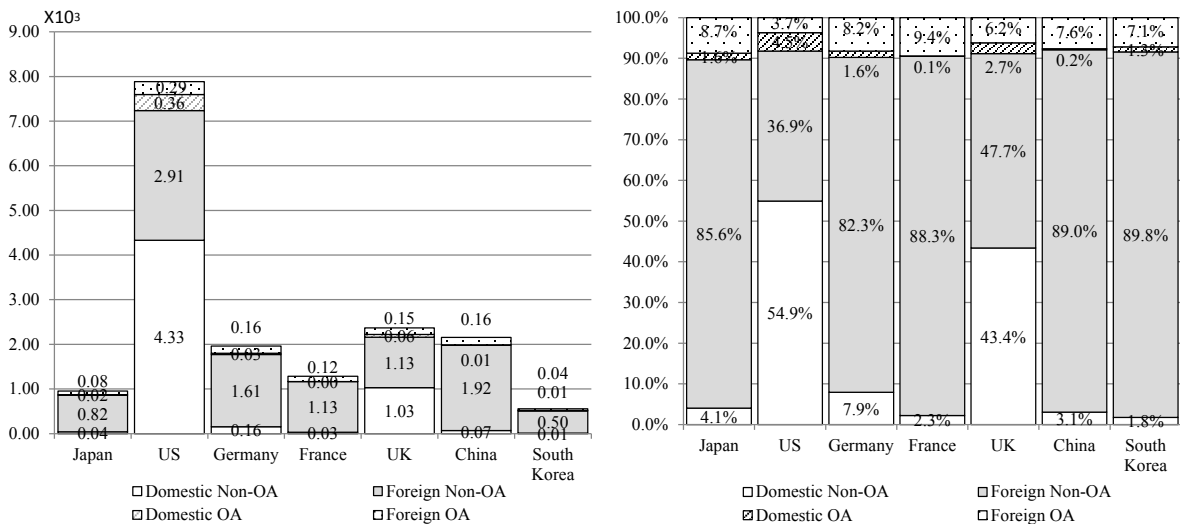


Figure 3: Mean number of papers and their ratio for the top 10% highly cited papers



Note: The top 10% highly cited papers in any of 27 fields are counted. The ratio of those to all papers does not necessarily become 10%.

Second, regarding the contribution ratio to the increase rate of the number of papers for the two periods (2004-06, 2010-12) (Table 1), Japan had a unique trend where an increase in the number of foreign OA journals' papers contributed to an increase in the number of papers. In other countries, an increase in the number of foreign Non-OA journals' papers contributed to an increase in the number of papers.

Table 1: The contribution ratio to the increase rate of the number of papers for the two periods

All Papers	The increase rate of the no. papers for the two periods	Contribution ratio			
		Domestic Non-OA	Foreign Non-OA	Domestic OA	Foreign OA
Japan	4.3%	1.0%	-2.6%	0.6%	5.3%
US	21.7%	3.9%	10.7%	3.5%	3.6%
Germany	26.4%	1.4%	16.1%	1.2%	7.7%
France	27.0%	1.0%	17.9%	0.2%	8.0%
UK	27.7%	10.2%	9.4%	2.7%	5.3%
China	91.9%	19.1%	61.7%	1.5%	9.6%
South Korea	96.1%	15.7%	61.9%	7.7%	10.7%

Top10% highly cited papers	The increase rate of the no. papers for the two periods	Contribution ratio			
		Domestic Non-OA	Foreign Non-OA	Domestic OA	Foreign OA
Japan	0.9%	-0.4%	-5.0%	0.4%	5.9%
US	11.5%	-3.2%	8.8%	3.3%	2.6%
Germany	37.9%	2.9%	25.4%	1.7%	7.9%
France	33.7%	0.1%	24.3%	-0.1%	9.4%
UK	28.9%	10.3%	10.7%	1.9%	6.0%
China	167.0%	4.2%	144.2%	0.3%	18.2%
South Korea	75.1%	2.3%	61.2%	1.7%	9.9%

Third, if publications are made in English, OA journals' papers get the opportunity of being cited in other countries. Table 2 shows the ratio of number of papers by language used in the domestic journals⁵, where I analyzed the language of the main text. If not OA, in countries other than France and China, more than half of the papers were in English. However, for the top 10% highly cited papers, in countries other than China, English was used for almost all papers. In the case of OA journals, almost all countries except France and China use English, and that tendency became stronger for the top 10% highly cited papers.

Table 2: Ratio of number of papers by language used in the domestic journals

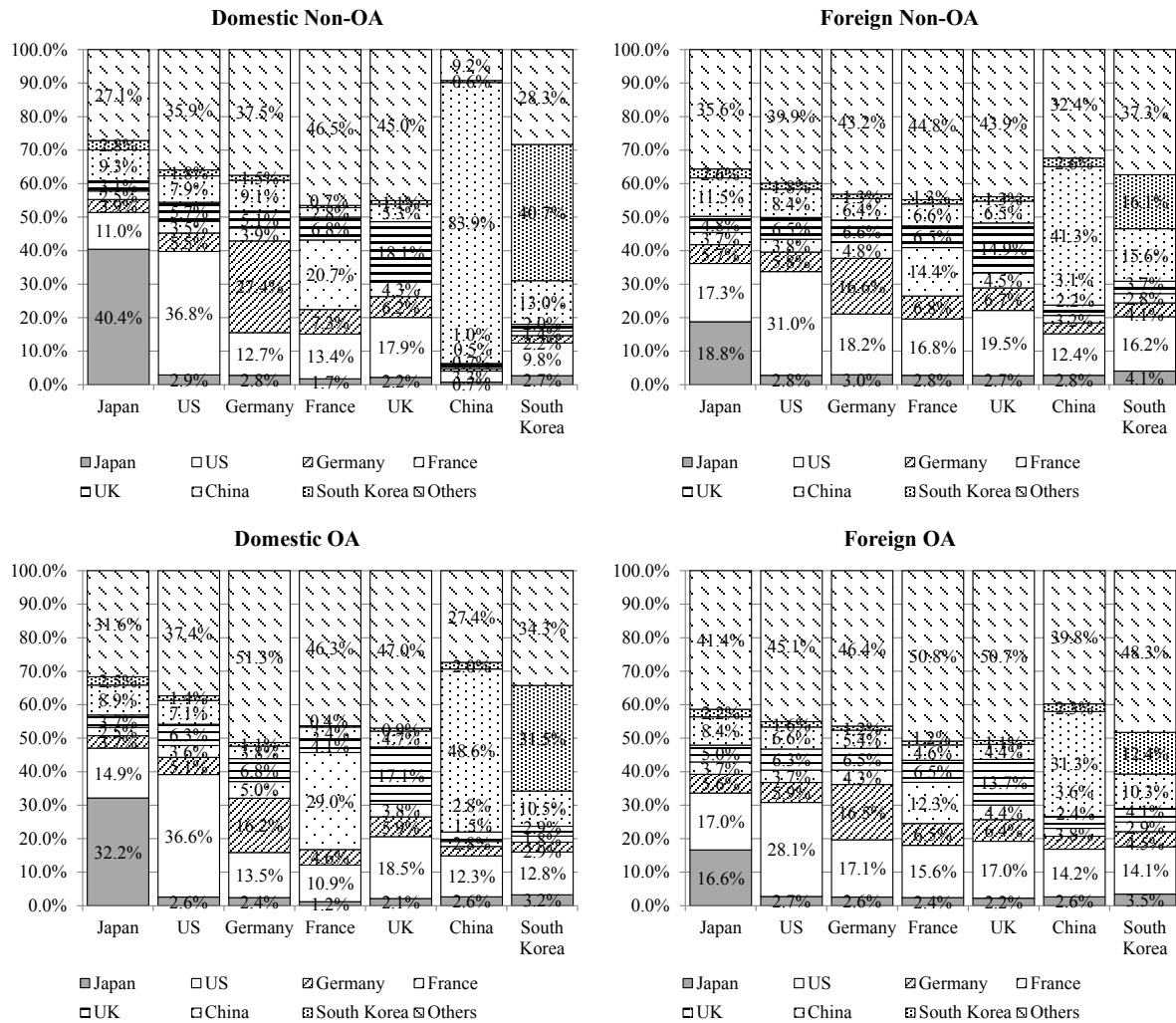
Country	Domestic journal							
	Non-OA				OA			
	All papers		Top10% papers		All papers		Top10% papers	
	Native	English	Native	English	Native	English	Native	English
Japan(Japanese)	46.5%	53.2%	0.9%	99.1%	10.2%	89.6%	0.0%	100.0%
US(English)	-	99.8%	-	99.9%	-	99.9%	-	99.9%
Germany(Germany)	50.2%	49.7%	4.0%	96.0%	9.8%	90.1%	4.8%	95.2%
France(French)	77.6%	22.0%	8.1%	91.8%	73.9%	25.4%	47.6%	52.4%
UK(English)	-	99.8%	-	99.9%	-	100.0%	-	100.0%
China(Chinese)	86.8%	13.2%	64.9%	35.1%	42.3%	57.1%	3.8%	96.2%
South Korea(Korean)	29.3%	70.7%	9.1%	90.9%	7.7%	92.3%	0.0%	100.0%

Note: The native language is written in parentheses.

Finally, Figure 4 shows the ratio of citing countries of each journal type's papers to identify knowledge diffusion among counties. OA journal's papers decrease the ratio of own country citations.

⁵ For foreign journals, English is used in approximately 95% or more of them in all countries.

Figure 4: The ratio of citing countries



DISCUSSION

This report revealed that the number of foreign Non-OA journal’s papers contributed to an increase in the number of papers except Japan. A certain amount of Non-OA journals’ paper was written in native language in the non-English-speaking countries. On the other hand, almost all OA papers written in English. The international collaboration ratio of OA journal’s papers is high. In addition, OA journals’ papers are better cited by foreign countries than Non-OA journals’ papers. It is suggested that when we consider OA policy, this related to language, international collaboration, and citation pattern.

REFERENCES

Sotudeh, H., Ghasempour, Z., and Yaghtin, M. (2015). The citation advantage of author-pays model: the case of Springer and Elsevier OA journals. *Scientometrics*, 104(2), 581-608.

Wang, X., Liu, C., Mao, W., and Fang, Z. (2015). The open access advantage considering citation, article usage and social media attention. *Scientometrics*, 103(2), 555-564.

Wang, L., Liu, X., and Fang, H. (2015). Investigation of the degree to which articles supported by research grants are published in open access health and life sciences journals. *Scientometrics*, 104(2), 511-528.



4D Specialty Approximation: Ability to Distinguish between Related Specialties

Nadine Rons*

*Nadine.Rons@vub.ac.be

Research Coordination Unit, Vrije Universiteit Brussel (VUB), Pleinlaan 2, B-1050 Brussels (Belgium)

ABSTRACT

Publication and citation patterns can vary significantly between related disciplines or more narrow specialties, even when sharing journals. Journal-based structures are therefore not accurate enough to approximate certain specialties, neither subject categories in global citation indices, nor cell sub-structures (Rons, 2012). This paper presents first test results of a new methodology that approximates the specialty of a highly specialized seed record by combining criteria for four publication metadata-fields, thereby broadly covering conceptual components defining disciplines and scholarly communication. To offer added value compared to journal-based structures, the methodology needs to generate sufficiently distinct results for seed directories in related specialties (sharing subject categories, cells, or even sources) with significantly different characteristics. This is tested successfully for the sub-domains of theoretical and experimental particle physics. In particular analyses of specialties with characteristics deviating from those of a broader discipline embedded in can benefit from an approach discerning down to specialty level. Such specialties are potentially present in all disciplines, for instance as cases of peripheral, emerging, frontier, or strategically prioritized research areas.

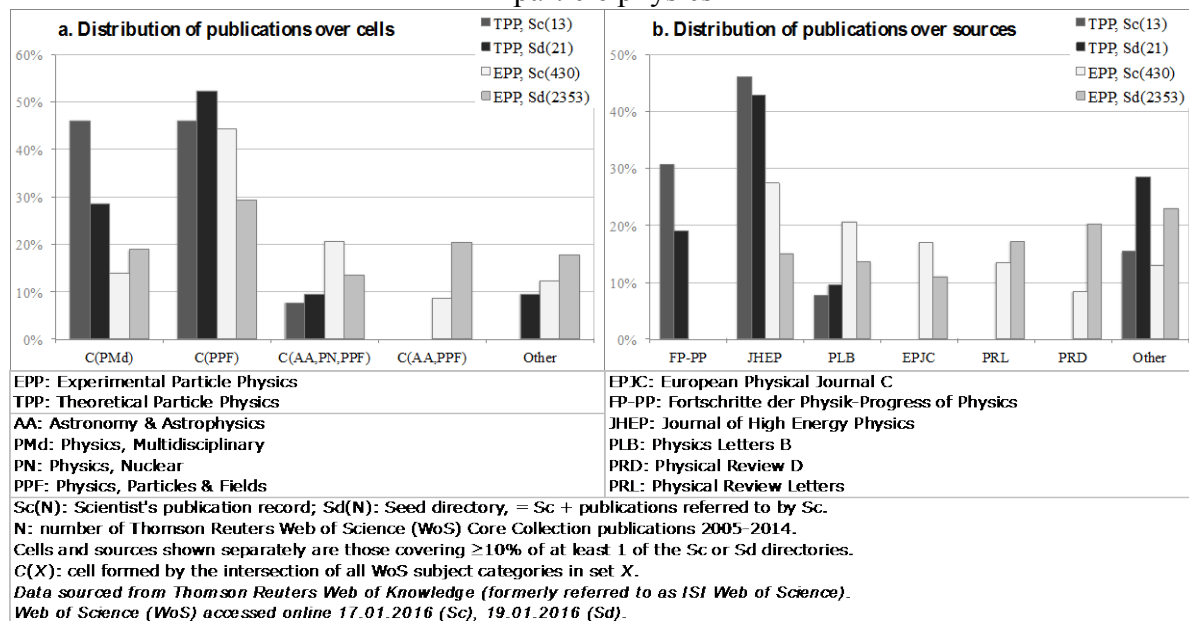
METHODOLOGY

The methodology uses four publication metadata-fields (references, authors, title, source) that are generally available in global citation indices such as Thomson Reuters' Web of Science (WoS) and Elsevier's Scopus. These are related to four of the six conceptual components defining disciplines as synthesized by Sugimoto and Weingart (2015) (cognitive, social, communicative, separatedness; the other two being: tradition, institutional) and to the four facets of the framework for bibliometric analysis of scholarly communication as proposed by Ni, Sugimoto and Cronin (2013) (artifact, producer, concept, gatekeeper). Combinations of some of these dimensions have been used previously to identify publication sets associated to particular research areas in various contexts (e.g. mapping, normalization, information retrieval). To the best of the author's knowledge, the proposed methodology is the first to bibliometrically approximate a specialty using criteria with this breadth of coverage of related conceptual components. It can be applied to publication records as specialized as those of an individual scientist or a team's research programme, provided that the seed directory is enlarged with publications referred to (diversifying authors while mainly adding publications in the same specialty, or at least in case of non-interdisciplinary research).

In a first phase, most frequently occurring 'key values' are selected in each dimension (references, authors, title words, cells) until a pre-set percentage (coverage threshold) of publications in the seed directory is covered by key values. In a second phase, the specialty is

approximated by the set of publications covered by key values in at least three of the four dimensions. In both phases coverage of a publication by key values in a particular dimension requires the publication to be associated to at least one key value for authors, references, and cells, and to at least two key values for title words. The possibility not to be associated to key values in one of the four dimensions prevents exclusions based on that dimension only of otherwise complying publications (false negatives). The required association to key values in at least three dimensions prevents inclusions based on one or two dimensions only of otherwise non-complying publications (false positives). The combination of dimensions also allows complexity per dimension to remain low. In the calculations for this paper, the coverage threshold was set to 80% for all dimensions, and key values were limited to words of at least five characters, to references in WoS identified via DOI, and to reprint authors (processed based on name and first initial, excluding frequently occurring names).

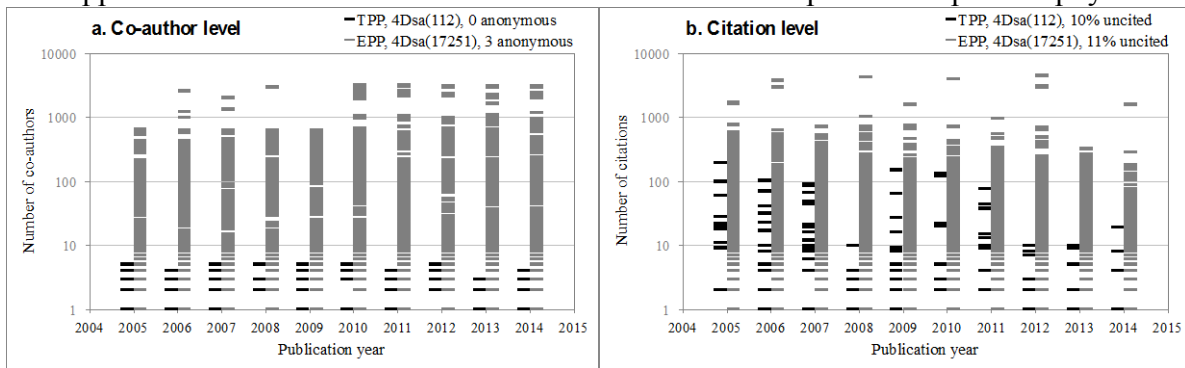
Figure 1. Proximity in publication venues of two team leaders in theoretical and experimental particle physics



RESULTS

Figure 1 illustrates the proximity in publication venues of two team leaders in theoretical and experimental particle physics, substantially sharing cells and even sources published in and referred to. These sub-domains are nevertheless known to strongly differ in attained numbers of co-authors and citations, much higher for experimental particle physics. In the related subject categories, cells, and even sources, these different cultures are blended. The scientists' 4D specialty approximations keep these traditions apart (no overlap), and strongly differ in attained levels of co-authors and citations (a few co-authors and several hundred citations for theoretical particle physics, versus several thousand for experimental particle physics), reflecting the known differences between these sub-domains (Figure 2).

Figure 2. Differences in attained co-author and citation levels between 4D specialty approximations for two team leaders in theoretical and experimental particle physics



EPP: Experimental Particle Physics; TPP: Theoretical Particle Physics.

4Dsa(N): 4D specialty approximation.

N: number of Thomson Reuters Web of Science (WoS) Core Collection publications 2005-2014.

Data sourced from Thomson Reuters Web of Knowledge (formerly referred to as ISI Web of Science).

Web of Science (WoS) accessed online 17.01.2016-07.03.2016.

DISCUSSION

Specialty approximations can provide information for various analyses with quantitative aims (e.g. reference values for normalized indicators, thresholds for outstanding performance) or qualitative aims (e.g. lists of potential peers, benchmarks, literature of interest). Whether a specialty approximation is sufficiently accurate (precise in delineation and complete in coverage) depends on the information to be derived from it, and on how strongly this information varies between related specialties. Also sub-specialties (e.g. dedicated to specific natural species or medical treatments) can have partly different inherent or contextual characteristics, resulting in different bibliometric characteristics. This paper demonstrates that the newly developed 4D specialty approximation methodology has the ability to generate distinct, coherent results from seed directories in closely related specialties, reflecting known differences in publication and citation characteristics. This requirement being met, a next step is to investigate the approximation's adequacy (inclusion of peers, confirmation by scientists, potential bias, ways to enhance specialty coverage or delineation precision). A sufficient level of confidence reached, utility as a basis for assessment, trend analysis, recommendation, benchmarking, and distinction between different kinds of research (basic/applied, theoretical/empirical, mono-disciplinary/interdisciplinary) are among potential paths for exploration.

REFERENCES

- Ni, C., Sugimoto, C.R., Cronin, B. (2013). Visualizing and comparing four facets of scholarly communication: Producers, artifacts, concepts, and gatekeepers. *Scientometrics*, 94(3), 1161-1173.
- Rons, N. (2012). Partition-based Field Normalization: An approach to highly specialized publication records. *Journal of Informetrics*, 6(1), 1-10.
- Sugimoto, C.R., Weingart, S. (2015). The kaleidoscope of disciplinarity. *Journal of Documentation*, 71(4), 775-794.



Analysis of Structure of Scientific Publications at Universities Focusing on Sub-Organizations

Akiyoshi MURAKAMI*, Ayaka SAKA* and Masatsura IGAMI*

*a-murakami@nistep.go.jp

Research Unit for Science and Technology Analysis and Indicators, National Institute of Science and Technology Policy (NISTEP), Kasumigaseki, Tokyo 100-0013 (JAPAN)

ABSTRACT

Inter-university competition is intensifying in various countries due to the progress of globalization. In Japan, functions required at universities are diversifying since the conversion of national universities to corporations in 2004, and university reform for each university to realize and enhance its own strengths and characteristics is required.

In the structure of scientific publications in Japan, about 70% of total publications in Japan are produced by higher education sector, indicative of universities playing a major role in the scientific knowledge creation in Japan.

The National Institute of Science and Technology Policy (NISTEP) has conducted analysis of each Japanese university with high-level research capabilities, and discovered that each university shows unique field composition in scientific publications [1]. However, it has not been identified from which sub-organizations of universities' scientific publications are produced for the field where the university has strengths in, and there has been no association made between the organization structure of universities and the situation of scientific publications.

In this study, we investigated the linkage between sub-organizations of universities and the fields of scientific publications for the purpose of analyzing the scientific publications structure of universities in detail.

METHODS

The NISTEP Dictionary of Names of Universities and Public Organizations [2] containing the name list of Japanese universities and public research institutions, etc. was used for name identification of the Web of Science (WoS) database of Thomson Reuters at sub-organization level.

In the dictionary, the name list is gathered for sub-organizations (lower organizations directly under universities) for universities with a certain scale, such as faculties, graduate schools, and attached institutes. The 32 universities (24 national, 2 public and 6 private) for which name identification was conducted for sub-organizations are listed in Table 1. For name identification, a program specifically developed by NISTEP was used.

The number of sub-organizations shown in Table 1 refers to those that appeared in scientific papers published during the period from 2009 to 2013. The identification rate refers to the rate of publications for which sub-organizations were identified among scientific papers published during the period from 2009 to 2013. The rate normally does not reach 100%, since some journals show only the names of universities in the affiliation of authors and some papers show names of departments that are at lower than the sub-organizations of universities. In the

present analysis, the mean identification rate of 31 universities is about 92% (except for the rate of 46% for the Tokyo Medical and Dental University).

Table 1. The 32 Universities Subject to Analysis

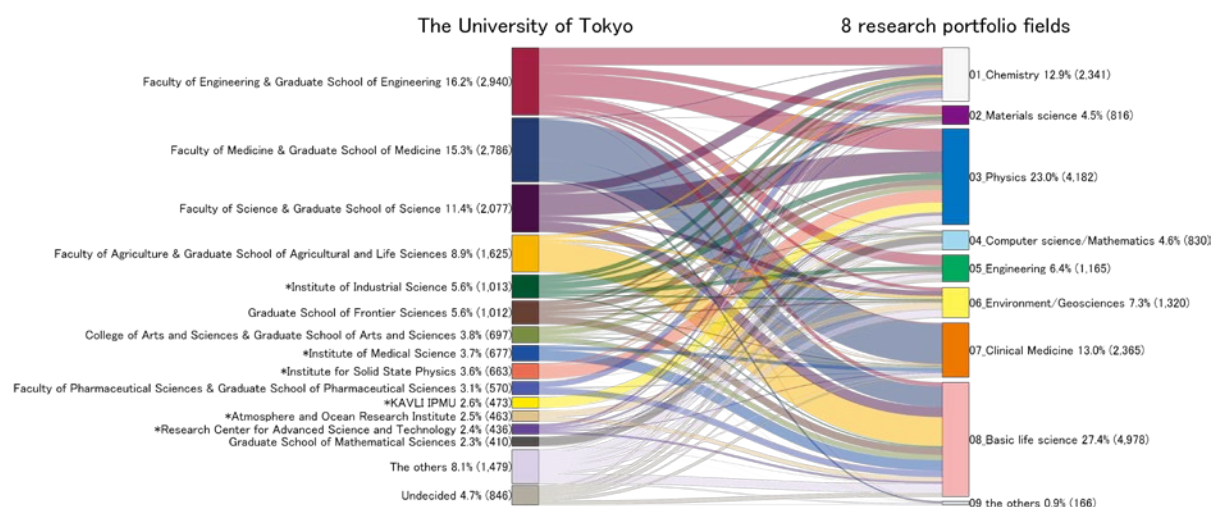
No	Name of university	Category	Number of sub-organizations	Identification rate
1	Chiba University	National	36	94%
2	Gifu University	National	22	94%
3	Gunma University	National	12	94%
4	Hiroshima University	National	39	92%
5	Hokkaido University	National	58	95%
6	Kanazawa University	National	28	94%
7	Keio University	Private	30	94%
8	Kinki University	Private	30	94%
9	Kobe University	National	32	95%
10	Kumamoto University	National	28	92%
11	Kyoto University	National	60	95%
12	Kyushu University	National	89	93%
13	Nagasaki University	National	24	92%
14	Nagoya University	National	50	94%
15	Nihon University	Private	33	96%
16	Niigata University	National	22	93%
17	Okayama University	National	31	93%
18	Osaka City University	Public	16	88%
19	Osaka Prefecture University	Public	19	89%
20	Osaka University	National	52	96%
21	Shinshu University	National	18	93%
22	The University of Tokyo	National	69	95%
23	Tohoku University	National	53	96%
24	Tokai University	Private	45	87%
25	Tokushima University	National	24	91%
26	Tokyo Institute of Technology	National	45	90%
27	Tokyo Medical and Dental University	National	17	46%
28	Tokyo University of Agriculture and Technology	National	15	83%
29	Tokyo University of Science	Private	20	91%
30	University of Toyama	National	23	90%
31	University of Tsukuba	National	65	88%
32	Waseda University	Private	52	81%

RESULTS

THE UNIVERSITY OF TOKYO

The linkage between the sub-organizations of the University of Tokyo (the university with the largest publications in Japan) and eight research portfolio fields is shown in Figure 1. Here, eight research portfolio fields refer to eight natural science fields compiled from the 22 fields of Essential Science Indicators (ESI) by Thomson Reuters. The number of papers is the total for five years from 2009 to 2013, and is counted by the fractional counting method based on affiliations.

From Figure 1, for the case of the University of Tokyo, it was found that the linkage-structure is highly complex, that is, each sub-organization of the university carries out research across multiple fields.

Figure 1: Linkage between sub-organizations of the University of Tokyo and eight research portfolio fields

Summarized by authors based on Thomson Reuters Web of Science XML (SCIE, 2014 year-end version). The types of papers subject to analysis were Articles and Reviews, and the fractional counting method was used. The number of papers is the total for five years from 2009 to 2013.

The sub-organizations (a sum of faculties and graduate schools) of the University of Tokyo with the largest share were the Faculty of Engineering & Graduate School of Engineering (16.2%), followed by the Faculty of Medicine & Graduate School of Medicine (15.3%), the Faculty of Science & Graduate School of Science (11.4%), and the Faculty of Agriculture & Graduate School of Agricultural and Life Sciences (8.9%).

The characteristic of the University of Tokyo is that research institutes which are shown by “*” in figure 1 have relatively large shares in addition to faculties and graduate schools, such as the Institute of Industrial Science (5.6%), Institute of Medical Science (3.7%), Institute for Solid State Physics (3.6%), KAVLI IPMU (2.6%), Atmosphere and Ocean Research Institute (2.5%), and Research Center for Advanced Science and Technology (2.4%).

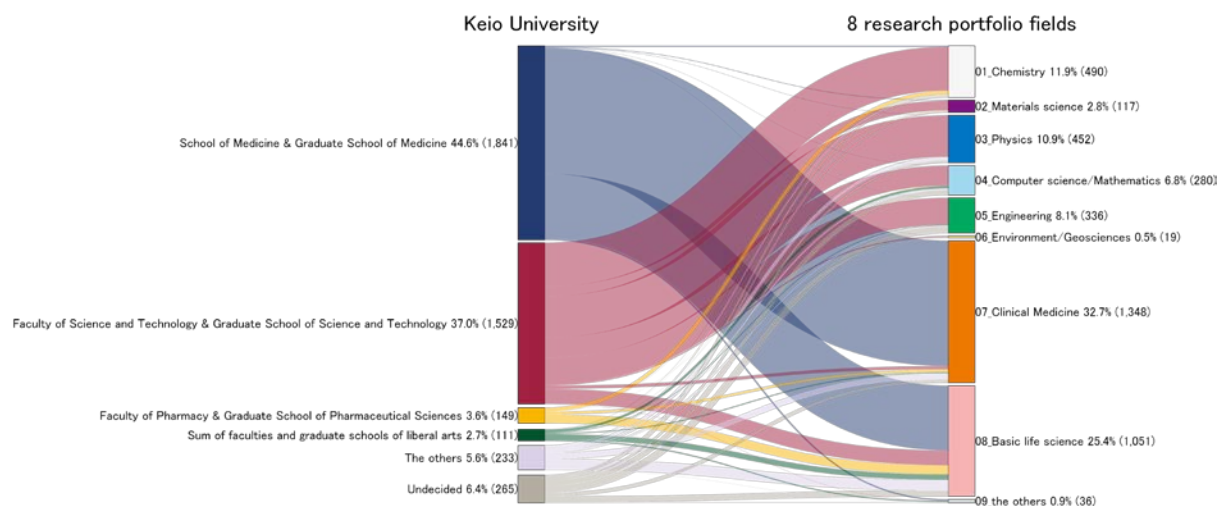
Interestingly, a large majority of publications produced by the Faculty of Engineering & Graduate School of Engineering were attributed to “01_Chemistry”, “03_Physics” and “02_Materials science”, and the percentage of “05_Engineering” which meets the name of the organization is relatively small. While the University of Tokyo has strong research capabilities in the field of physics [1], it was found that the research capabilities are comprised of multiple sub-organizations.

KEIO UNIVERSITY

The linkage between the sub-organizations of Keio University (the university with the largest publications scale among private universities in Japan) and eight research portfolio fields is shown in Figure 2. At Keio University, the sub-organizations (a sum of faculties and graduate schools) with the largest share were the School of Medicine & Graduate School of Medicine (44.6%), followed by the Faculty of Science and Technology & Graduate School of Science and Technology (37.0%) and the Faculty of Pharmacy & Graduate School of Pharmaceutical Sciences (3.6%). While a large majority of publications by the School of Medicine & Graduate School of Medicine were attributed to 2 fields, namely, “07_Clinical Medicine” and “08_Basic life science”, publications by the Faculty of Science and Technology & Graduate School of Science and Technology were distributed to six fields, namely, “01_Chemistry”, “02_Materials Science”, “03_Physics”, “04_Computer science/Mathematics”,

“05_Engineering” and “08_Basic life science”. Keio University has strengths in clinical medicine and basic life science [1], and it was found that the strengths are attributable to the School of Medicine & Graduate School of Medicine.

Figure 2: Linkage between sub-organizations of Keio University and eight research portfolio fields



Summarized by authors based on Thomson Reuters Web of Science XML (SCIE, 2014 year-end version). The types of papers subject to analysis were Articles and Reviews, and the fractional counting method was used. The number of papers is the total for 5 years from 2009 to 2013.

DISCUSSION

In the present research, the sub-organization level analyses of scientific publications were made, and its linkage with research fields was shown for the examples of the University of Tokyo and Keio University. From the comparison between the University of Tokyo and Keio University, it was clarified that the former has a complicated linkage-structure where multiple sub-organizations produce publications and each of them carries out research in multiple fields, while the latter has a simple linkage-structure where a small number of organizations produce a large majority of publications. This result implies that organization management could be different when universities seek to have the strength in a specific research field depending on the linkage-structure.

In the future, further investigation will be carried out for universities other than the two universities above to analyze the scientific publications structure of each university at sub-organization level, which may serve as useful evidence for science policy and university management.

REFERENCES

- [1] Saka, A. & Igami, M. (2015). Benchmarking Research Capability of Japanese Universities 2015, Research Material No. 243, NISTEP (*in Japanese*)
- [2] Onodera, N., Igami, M. & Saka, A. (2015). Preparation and Utilization of the NISTEP Dictionary of Names of Universities and Public Organizations, NISTEP NOTE No. 15, NISTEP (*in Japanese*)

INDEX BY AUTHOR

AUTHOR	PAGE
Abad-García, María F.	419
Adams, Jonathan	332
Aduku, Kuku J.	1165
Aguillo, Isidro	1147
Ahn, Yong-Yeol	1160
Aksnes, Dag W.	1214
Al, Umut	1494
Albarrán, Pedro	1241
Aleixandre Benavent, Rafael	1315, 1487
Aleixandre Mendizábal, Guillermo	246
Aleskerov, Fuad T.	1266
Alfonso, William H.	996
Al-Hashmi, Shatha	272
Alkærsig, Lars	1076
Almeida, Catia C.	1491
Alperin, Juan Pablo	58, 477
Al-Saleh, Yasser	272
Álvarez-Bornstein, Belén	1345
Aman, Valeria	1042
Amaradasa, Ranasinghe	1059
Amaro, Marcela	1016
Amigó, Vicente	1292
Andersen, Jens Peter	733, 852
Aouinaït, Camille	497
Archambault, Éric	832, 1229
Archambeault, Jean	289
Arias, Begoña	981
Arora, Lucy	832
Åström, Fredrik	462
Azagra-Caro, Joaquín M.	213, 1370
Baard, Patrik	780
Baranova, Olga	1433
Barge-Gil, Andres	208
Bar-Ilan, Judit	1156
Barjak, Franz	562
Barrenechea, Julieta	971
Barreto, Jorge	1325
Barreto, Maurício	1325
Barriere, Stina	780
Beaudry, Catherine	718, 819
Belda-Miquel, Sergio	981
Bello, Alessandro	808
Benneworth, Paul	511
Bergé, Laurent	1106

AUTHOR	PAGE
Bianco, Mariela	313
Blasco, Agustín	1292
Bloch, Carter	733
Boavida, Nuno	156
Boni, Alejandra	981
Bordons, María	1345
Bornmann, Lutz	1200
Boshoff, Nelius	325
Botti, Vicent	1292
Bowman, Sarah	1190
Bowman, Timothy D.	477, 1190
Boyack, Kevin W.	523
Bozeman, Barry	617
Brás, Oriana Rainho	406
Bretschneider, Stuart	617
Brozaitis, Haroldas	388
Buchnik, Tsipy	1355
Buehrer, Susanne	375
Bugge, Markus	222
Buhmann, Alexander	904
Cabanac, Guillaume	812
Cai, Sanfa	1471
Calero-Medina, Clara M.	73, 90
Caloghirou, Yannis	192, 197
Callaert, Julie	832
Cambrosio, Alberto	406
Cameron-Pesant, Sarah	1435
Campbell, David	832, 1229
Cano, Juan	1292
Cañibano, Carolina	644, 771
Capilla, José	1292
Capmany, José	1292
Carlen, Christoph	497
Carley, Stephen	622
Caron, Emiel	1405
Castanha, Renata	1482
Castelló-Cogollos, Lourdes	1315, 1487
Castelló-Cogollos, Rafael	1487
Castro, Nayibe	283
Castro-Martínez, Elena	511
Chavarro, Diego	66
Chen, Carey Ming-Li	1344
Chi, Pei-Shan	708, 1223
Chinchilla-Rodríguez, Zaida	662, 1412

AUTHOR	PAGE
Chiralt, Amparo	1292
Christen, Danilo	497
Conejero, J. Alberto	1292
Costas, Rodrigo	90, 598, 644, 1173, 1196
Côté, Grégoire	1229
Crespi, Gustavo	187
Cuesta, Salomé	1292
d'Este, Pablo	1315
David, Leonor	406
De Filippo, Daniela	303, 1033, 1454
De La Fuente, Teresa	981
de Rijcke, Sarah	516
Debacker, Noëmi	627
Dehdarirad, Tahereh	1463
Delong, Tang	1274
Denia, Elena	1301
Derrick, Gemma	318, 379, 516
Desrochers, Nadine	477, 598, 887
Díaz-Faes, Adrián A.	477, 598
Didegah, Fereshteh	1190
Ding, Kun	1395
Dobrescu, Simona	1011
Doğan, Güleda	1494
Donner, Paul	85
Duarte, Kedma	1310
Dyachenko, Ekaterina	1419
Eckert, Denis	671
Edquist, Charles	149
Efimenko, Irina	401
Engels, Tim C.E.	893
Esterhuyse, Harrie	325
Fang, Zhichao	1360
Feinstein, Osvaldo N.	357
Feltrin, Rebeca	1325
Fernández, Ernesto	808
Fernández-Mesa, Anabel	1370
Ferreira, Manuel E.	635
Fonseca e Fonseca, Bruna	442, 1325
Foray, Dominique	497
Fu, Huang	1274
Fukuzawa, Naomi	1506
Fursov, Konstantin	1005
Galsworthy, Mike	388
Galleron, Ioana	901
García Álvarez-Coque, José María	1292
Garechana, Gaizka	1384
Garnett, Alex	58
Gauffriau, Marianne	1467

AUTHOR	PAGE
Gemeinholzer, Birgit	97
Getz, Daphne	208, 1355
Ghiasi, Gita	844
Giménez-Toledo, Elea	940
Glänzel, Wolfgang	708, 1223
Gläser, Jochen	518
Gök, Abdullah	702
Gómez, Aitor	884
González-Alcaide, Gregorio	1440, 1444
Gorraiz, Juan	708
Gorry, Philippe	424
Gou, Xiaoli	1081
Gouveia, Fabio C.	1123
Grácio, Maria C. C.	1282, 1475, 1482, 1491
Griessler, Erich	375
Grossetti, Michel	671
Gumpenberger, Christian	708
Guner, Stanisław	1405
Guns, Raf	893, 957
Guoqiang, Liang	1274
Haddad, Carolina R.	139
Hardeman, Sjoerd	168
Hartley, James	1190
Haunschild, Robin	1200
Haustein, Stefanie	410, 477, 1196
Herrero-Solana, Víctor	1349
Heuritsch, Julia	741
Heyman, Ulf	531
Hicks, Diana	343
Hinze, Sybille	771, 1250
Hollanders, Hugo	127
Holt, Ilkay	439
Holzinger, Florian	863
Hopkins, Michael	485
Horellou, Süzal	575
Hornbostel, Stefan	771
Horstmann, Wolfram	1429
Hu, Xiaojun	1117
Hu, Zhigang	1274, 1360, 1395
Huamani, Charles	1440
Huang, Lu	252
Huang, Ying	1374, 1389
Hug, Sven E.	873
Ibarra, Andoni	971
Idrissou, Al	106
Igami, Masatsura	1514
Iizuka, Michiko	127
Ingenhoff, Diana	904

AUTHOR	PAGE
Isett, Kim	343
Iversen, Eric	771
Jacobsson, Carl	780
Jansen, Yorrick	1435
Jégou, Laurent	671
Jiménez-Contreras, Evaristo	1129
Joanpere, Mar	884
Johann, David	1332
Jonkers, Koen	607
Jordá, Carmen	1292
Joyal, Bruno	448
Juliá, Juan	1292
Julien, Charles-Antoine	797
Jung, Nicole	1304
Jürgens, Björn	1349
Kahn, Michael	653
Karaulova, Maria	702
Karlsson, Staffan	780
Kawamura, Takahiro	1449
Kawashima, Hirotaka	789
Ke, Qing	1160
Keating, Tamara	289
Khalili, Ali	106
Khoroshevsky, Vladimir	401
Kim, Sun-Woo	1478
Klavans, Richard	523
Kobashi, Nair Y.	1329
Kokorevics, Arnis	1318
Kostinskiy, Alexander	550
Kousha, Kayvan	1165
Kramer, Claudia	1304
Lalonde, Bastien St. Louis	832
Lamers, Wout S.	1498
Lang, Frédérique	485
Laranja, Manuel	156
Larédo, Philippe	575
Larivière, Vincent	410, 448, 477, 598, 766, 797, 844, 887, 1196, 1435
Larruscain-Sarasola, Jaso	1384
Laurens, Patricia	575
Lavik, Gry Ane	934
Lee, June Young	1478
Leivas, Monique	981
Lepori, Benedetto	497, 607, 904
Leta, Jacqueline	812
Levitt, Jonathan	453
Leydesdorff, Loet	1117
Lima, Nísia Trindade	1325

AUTHOR	PAGE
Lindahl, Jonas	759
Lindner, Ralf	375
Liu, Chichen	1471
Liu, Yuxian	1471, 1502
Llorente, Pedro	1444
Loach, Tamar	332
Loizou, Antonis	106
López Cerezo, José Antonio	1012
López López, Wilson	1147
López-Fogués , Aurora	981
Luan, Chunjuan	1050
Lucio-Arias, Diana	283
Luwel, Marc	682
Lv, Qi	1374, 1389
Macaluso, Benoît	598, 766
Mahrholz, Nadine	115
Mahroum, Sami	269, 272
Maisonobe, Marion	671
Maital, Shlomo	1355
Makkonen, Teemu	1025
Malgarini, Marco	915
Manola, Natalia	388
Mantouvalou, Katerina	832
Mañana-Rodríguez, Jorge	940
Maroulis, Nikos	375
Marrades, Ramon	1011
Martens, Silvia	881
Martins, Fatima	1325
Matsumura, Katsuji	1449
Mauleón, Elba	1454
Mayer, Sabrina	924, 1458
McLean, Robert K.D.	357
Meijer , Ingeborg	375, 419, 516, 803
Mejlgaard, Niels	375
Melkers, Julia	343, 571
Mendonça, Sandro	635
Metaxas, Omiros	388
Miguel, Pedro	1292
Miguel, Sandra	662, 1412
Mika, Pavel	78
Mikki, Susanne	1298
Milard, Béatrice	671
Minkoua, Jules René	366
Miranda, Miguel	1292
Mirnezami, Seyed Reza	718
Mitkova, Liliana	252, 1374
Modrego Rico, Aurelia	208
Moed, Henk	693

AUTHOR	PAGE
Molas-Gallart, Jordi	516
Mongeon, Philippe	410, 448, 477
Moniz, António	156
Mora Holguín, Henry	283
Morales, Alberto	1016
Mori, Andres	1426
Morillo, Fernanda	1345
Muellerleile, Christopher	51
Mugnaini, Rogério	419
Muñoz van den Eynde, Ana	1012
Murakami, Akiyoshi	1514
Nefedova, Alena	1005
Nelhans, Gustaf	393
Nepelski, Daniel	555
Neufeld, Jörg	1332
Newfield, Christopher	51
Nielsen , Mathias Wullum	852
Nochevkin, Evgeny	401
Noyons, Ed	401
Nunes, João A.	406
Nuredini, Kaltrina	1137
Ochsner, Michael	873
Ogsuz, Gizem	1374
Okamura, Asako	990
Olechnicka, Agnieszka	1208
Oliveira, Ely F. T.	1282
Olmeda-Gómez , Carlos	662, 1412
Olmos-Peñuela, Julia	511
O'Neill, Eoghan	555
Ovalle-Perandones, María-Antonia	662, 1412
Pacheco, Roberto C. S.	1310
Pajevic, Filipa	1011
Pálinkó, Éva	750
Pandiella-Dominique, Andres	303
Pardo Martínez, Clara I.	283, 996
Park, Jinseo	1440, 1478
Paul-Hus, Adèle	477, 598
Pellicer-Sifres, Victoria	981
Peng, Ming-Te	1307
Peraza Castaneda, Elias Humberto	246
Pereira, João	635
Perianes-Rodriguez, Antonio	662, 1173, 1257, 1412
Peris, Alfred	1292, 1433
Pesole, Annarosa	555
Peter, Viola	375
Peters, Isabella	1137, 1156
Petrie, Stephen	298
Piro, Fredrik N.	1214

AUTHOR	PAGE
Pislyakov, Vladimir	1266
Plonski, Guilherme A.	208
Ploszaj, Adam	1208
Pooladian, Aida	1463
Porter, Alan L.	622, 1389
Prins, Ad	965
Protogerou, Aimilia	197
Quan-Haase, Anabel	477
Rafols, Ismael	66, 347, 375, 419, 485, 1183, 1315
Ramos, José M.	1440, 1444
Ravn, Tine	375
Reale, Emanuela	590, 607, 803
Reinhardt, Anke	122
Reinhold, Anke	115
Reis, Juliana Gonçalves	1325
Repiso, Rafael	1095
Rio Belver, Rosa María	1384
Rittberger, Marc	115
Roberge, Guillaume	1229
Robinson, Douglas K.R.	575
Robinson-Garcia, Nicolas	644, 1095, 1129, 1183, 1370
Rons, Nadine	1511
Rørstad, Kristoffer	1214
Rostovtsev, Andrey	550
Rotolo, Daniele	1089
Rousseau, Ronald	1117
Ruiz-Castillo, Javier	1173, 1241, 1257
Ryan, Thomas K.	581, 733
Sainte-Marie, Maxime	598
Saka, Ayaka	1514
Salazar, Monica	187
Samano , Elvira	884
Sampaio, Ricardo B.	442, 1325
Samuel, Gabrielle	318, 379
Sánchez-Ruiz, Luis	1292
Sandström, Ulf	531
Santos, Paula X.	442, 1325
Sanz-Casado, Elías	1033
Schaaper, Martin	808
Schatz, Wolfgang	881
Scherngell, Thomas	1106
Schiffbaenker, Helene	863
Schlobach, Stefan	106
Schmidt, Birgit	97, 1429
Schmidt, Nora	951
Schneider, Jesper W.	581, 852
Schoen, Antoine	575
Shang, Lining	252

AUTHOR	PAGE
Shapira, Philip	702
Shu, Fei	410, 797
Silva, Deise	1475
Silva, Marcus V.	442
Siokas, Evangelos	197
Sivertsen, Gunnar	78, 901, 910, 934
Smith, Elise	410, 448, 477
Solberg, Espen	222
Solomon, Gregg	622
Song, Le	343
Song, Tae Ho	1478
Sotudeh, Hajar	1463
Souza, Claudia	1033
Souza, Iara V.P.	1123
Spaapen, Jack	965
Spinello, Andrea	590
Stahlschmidt, Stephan	1250
Stanciauskas, Vilius	388
Stilgoe, Jack	375
Stout, Luka	1081
St-Pierre, Carl	819
Stranack, Kevin	58
Subochev, Andrey N.	1266
Sugimoto, Cassidy R.	477, 766, 844, 1160
Sun, Xiaoling	1395
Sutz, Judith	313
Szarzec, Jakub	78
Szomszor, Martin	332, 1426
Talón, Alba	981
Tang, David L.	289
Tang, Puay	66
Taşkın, Zehra	1494
Temple, Ludovic	366
Tesch, Jakob	771
Thelwall, Mike	453, 1165
Thurner, Thomas	1005
Tijssen, Robert J.W.	213, 1069
Tippett, Chantale	832, 1229
Tönurist, Piret	1339
Torres-Salinas, Daniel	1095, 1129
Tremouilhac, Pierre	1304
Treolar, Andrew	97
Trivedi, Rakshit	343
Tsipouri, Lena	375
Tumpane, John	780
Turpin, Tim	1059
Uribe-Tirado, Alejandro	1147
Uriona Maldonado, Mauricio	139

AUTHOR	PAGE
Valdmaa, Kaija	1339
Valente, Marco	467
Van den Besselaar, Peter	106, 531, 1081
van der Weijden, Inge C. M.	516, 741, 803
Van der Wurff, Lambertus J.	682
van Eck, Nees Jan	541
van Harmelen, Frank	106
van Honk, Jeroen	90
van Leeuwen, Thed	1183
van Raan, Anthony F.J.	1069
Van Roy, Vincent	555
van Vree, Frank	965
van Wijk, Erik	682
Vandevelde, Karen	627
Vargas, Fernando	234
Vértesy, Daniel	168, 229
Vienni, Bianca	506
Villavicencio, Daniel	1016
Vilsmaier, Ulli	506
Voegtler, Eva	637
Vogel, Michely J.M.	1329
Vonortas, Nicholas	192
Waaier, Cathelijn J. F.	741, 766
Waltman, Ludo	541
Wang, Hongyin	1360
Wang, Kangrui	252
Wang, Lili	178, 682
Wang, Xianwen	178, 1360
Wang, Xuefeng	1374, 1389
Wanzenböck, Iris	1106
Weber, Rosina	1310
Wei, Wenjie	1502
White, Lexi C.	435
Wilhelmsen, Lars	222
Wilsdon, James	1156
Windzio, Michael	637
Winnink, Jos J.	1069
Wiseman, Erica	289
Woltmann, Sabrina	1076
Woolley, Richard	375, 644, 771
Wouters, Paul	516, 1156
Wroblewski, Angela	375
Xianwen, Wang	1274
Yamashita, Yasuhiro	789, 1449
Yan, Erjia	100
Yang, Yang	1360, 1395
Yao, Junlan	1502
Yegros-Yegros, Alfredo	213, 419

AUTHOR	PAGE
Youtie, Jan	617, 622
Yu, Huizhu	252
Zabala-Iturriagagoitia, Jon Mikel	149
Zacharewicz, Thomas	607
Zahedi, Zohreh	1196
Zárate, Sandra	283
Zhang, Chunbo	1395

AUTHOR	PAGE
Zhao, Kangran	1405
Zhao, Mengnan	100
Zhigang, Hu	1274
Zhu, Donghua	1374, 1389
Zinilli, Antonio	590
Zuccala, Alesia	1095
Zygmuntowska, Marta	1298

ORGANIZERS



PLATINUM SPONSORS



UNIVERSITAT
POLITÀCNICA
DE VALÈNCIA

GOLD SPONSOR



THOMSON REUTERS®



Cátedra CTS+I

