Peer Review vs. Metrics - Abstracts

Professor Loet Leydesdorff

Abstract: “Ambitious People know their H-Factor.”

“Peer Review vs. Metrics” is no longer a valid opposition: the peers know each other’s h-index. The sciences can be considered as reputationally controlled work organizations. However, the mass production of scientific results makes it necessary to use also quantitative measures for quality control. In response, “citizen scientometrics”—e.g., knowing one’s h-index; using Google Scholar—has emerged alongside “professional” metrics. Professional metrics has become a (quasi-)industry.

Both peer review and metrics are affected by ongoing processes of formative meta-evaluation; both have pros and cons. Peer review has been bestowed with the status of “gold standard,” but peer review has also been an elitist and conservative control mechanism which tends to be blind for the error it generates. When one has to choose among the top-candidates in a final round, for example, peer review no longer works: irrational preferences enter the decision-making process. I argue for casting lots among top candidates in the final round.

Scientometric indicators reduce the complexity using a model. Models can always be improved; improving the model may reduce error. Large-scale models such as university rankings, however, are difficult to validate. Correlations among rankings can be low. Furthermore, the indicator industry has its own momentum. For example, scientometric evaluations continue to be benchmarked against the mean of a heavily skewed distribution (e.g., the “mean-normalized citation score”). Classifications generate serious problems at the borders between disciplines. The effects of the modelling on the results can sometimes be estimated.

The trade-off between peer review and metrics can be made the subject of (formative!) meta-evaluation; but the choice remains a *politicum*.

Bio:

Loet Leydesdorff (Ph.D. Sociology, M.A. Philosophy, and M.Sc. Biochemistry) is Professor at the Amsterdam School of Communications Research (ASCoR) of the University of Amsterdam. He is Associate Faculty at the Science and Technology Policy Research Unit (SPRU) of the University of Sussex, Visiting Professor of the Institute of Scientific and Technical Information of China (ISTIC) in Beijing, Guest Professor at Zhejiang University in Hangzhou, and Visiting Fellow at the School of Management, Birkbeck, University of London. He has published extensively in systems theory, social network analysis, scientometrics, and the sociology of innovation (see at http://www.leydesdorff.net/list.htm). With Henry Etzkowitz, he initiated a series of workshops, conferences, and special issues about the Triple Helix of University-Industry-Government Relations. He received the Derek de Solla Price Award for Scientometrics and Informetrics in 2003 and held “The City of Lausanne” Honor Chair at the School of Economics, Université de Lausanne, in 2005. In 2007, he was Vice-President of the 8th International Conference on Computing Anticipatory Systems (CASYS’07, Liège). Since 2014, Thomson Reuters lists him as a highly-cited author (http://highlycited.com).
ORCID: http://orcid.org/0000-0002-7835-3098
Researcher ID: E-2903-2010

loet@leydesdorff.net
http://www.leydesdorff.net

Professor Paul Edwards

Abstract: “Peer review: Still the Least Worst Measure of Research Performance.”

A view of metrics depends on what you are trying to measure. Overall indices such as citation counts can be useful indicators of the activity of a department or research group. The Business and Management REF sub-panel used data on such things as research income and PhD completions, but it interrogated them very closely. In its assessment of outputs it relied on reading the outputs, deploying several mean to calibrate individual sub-panel members' ratings. The aim of the REF was of course to give an overall score for a unit of assessment, not to rate individuals. The more that the focus is on individual performance, the more that [a] peer review of outputs and [b] other measures of scholarly contribution need to continue to be central to an assessment.

Bio:

Paul Edwards is a part-time professor of employment relations at Birmingham Business School, where he was previously head of the Department of Management. He was formerly at Warwick Business School, where among other roles he was Associate Dean for Research. He is a Fellow of the British Academy, and has been on 2 RAE/REF sub-panels for business and management. He is former editor of Work, Employment and Society and of Human Relations. Current research interests include the relationship between critical social science and public engagement.

Professor Eleonora Belfiore

Abstract: “Metrics vs. Peer Review: Time to Bury the Hatchet?”

The extent to which metrics can effectively lead to the assessment of research quality in academic publications is a longstanding matter of argument and debate. Attempts to shift quality assessment procedures to metrics-based approaches are usually resisted – by some academic communities more forcefully than others – in favour of peer review as the flawed, yet preferable means to judge the originality, soundness and intellectual contribution of academic scholarship.

The call for the use of ‘responsible metrics’ as support for peer review in the 2015 report *The Metrics Tide*, which I was involved in producing, is a testament to the complexity of the debate, and the likely impossibility (at least so far) to develop a faultless system to identify the best research across disciplinary boundaries. As a researcher working interdisciplinary at the intersection of the arts and humanities and social sciences, I will present the case for the importance of peer review in research quality assessment. However, this position acknowledges that peer review is far from a faultless system, and some of its weaknesses - in relation to matters of equality and gender, for example - are often not too dissimilar to those identified with metrics, hence my belief that the call for balance and the combination of diverse strategies for quality assessments that we made in *The Metrics Tide* remains, at present, the best way forward.

Bio:

Eleonora Belfiore is Professor of Communication and Media Studies at Loughborough University. She has published extensively on cultural politics and policy, and particularly the place that notions of the ‘social impacts’ of the arts have had in British cultural policy discourses. For Palgrave she has published, with Oliver Bennett, The Social Impact of the Arts: An intellectual history (2008) and co-edited with Anna Upchurch a volume entitled Humanities in the Twenty-First Century: Beyond Utility and Markets (2013). More recently, her research has focused on researching the politics of cultural value, and she was Director of Studies of the Warwick Commission on the Future of Cultural Value (2013-5), and co-author of its final report, Enriching Britain: Culture, creativity and growth, published in February 2015. Eleonora is currently developing new research on the cultural value of everyday forms of cultural participation in her role as co-investigator on the AHRC funded Connected Communities project ‘Understanding Everyday Participation – Articulating Cultural Values’. She was the lead for Loughborough University involvement as Founding Associate in the Tate Exchange initiative with Tate Museums in 2016-7.

Dr Liz Allen

Abstract: "Time for a more Grown-up Debate about Research Indicators."

The research community (including all players from researchers to institutions and information providers) currently get lost in endless debate about why quantitative indicators have limitations to support research evaluation – and equally, how peer review is both the 'gold standard' and flawed. Instead, we should harness this critique and dissatisfaction to try out new ways of describing and assessing research for it myriad of qualities – using and refining the indicators we currently have and thinking about new ones that can provide true insight into the questions we want to answer.

We want to do research efficiently. We want to incentivise researchers to do the best they can with the resources available. Funding agencies want to allocate their funding to help them to best achieve their objectives. It makes sense therefore to develop a more robust evidence base of what works and what might not– research on research / science of Science. A key part of this is thinking about how best to track, understand and assess research in all its contexts and uses.

Bio:

Liz Allen is Director of Strategic Initiatives at F1000, and involved in shaping new initiatives and partnerships to promote and foster open research. Prior to joining F1000 in 2015, Liz spent over a decade as Head of Evaluation at the Wellcome Trust, with a particular interest in impact assessment and the development of science-related indicators. In 2015 Liz became a Visiting Senior Research Fellow in the Policy Institute at King's College London. During 2014-15 Liz served as an adviser on the UK government commissioned Independent review of the role of research metrics in research assessment https://www.hefce.ac.uk/rsrch/metrics/. Liz was a Board Director of ORCID (www.orcid.org) from 2010 until 2015, and co-led the development of Project CRediT (Contributor Roles Taxonomy - http://www.casrai.org/CRediT).
Liz Allen - http://orcid.org/0000-0002-9298-3168

Professor John Mingers

Abstract: “Citations: Peer Review by the World.”

Research at all levels is increasingly subject to formal evaluation. In the UK we have suffered research assessment exercises over many years – indeed, they are probably the single biggest driver of university behaviour. So far, they have primarily been by peer review, said by many to be the “gold standard”. There have been some evaluations of the possibility of using metrics, as well or instead, including the recent “*Metrics Tide*” report but so far the use of metrics has been fairly marginal. In the 2014 REF, the Business and Management Panel actually said that it did not even want to see citation data.

On the other hand, there are many criticisms of peer review as being subjective, biased, and conservative as well as extremely expensive. At the same time, there are now huge amounts of citation data available in *WoS*, *Scopus* and *Google Scholar* and the discipline of scientometrics is generating ever more rigorous metrics and procedures. Is now not the time when metrics can be used in a major way to ease the burden and provide credible and reliable results that have less, or perhaps just different, biases and weaknesses?
In today’s workshop, we hope to have a lively and engaged debate between scholars from both sides and see whether there is, in fact, a way forward that can gain some degree of agreement.

For my part, I do see the value in metrics, which can be seen as peer review by the whole world, and I present some results based on Google Scholar data and pose the question; *does peer review actually produce any added value?*

Bio:

John Mingers is Professor of Operational Research and Systems at Kent Business School, and past Director of the School and Director of Research. His research interests include: research metrics and research performance; philosophy and information systems; critical realism; multimethodology - the combining of different methods within an intervention; autopoiesis; the nature of information, meaning and knowledge; and critical management studies. He is an Academician of the *Academy of the Social Sciences*, and has been on the Editorial Board of *MIS Quarterly, European Journal of Information Systems, Systems Research and Behavioural Science* and the *J. of Mixed Methods Research*. He published four books, the latest being *Systems Thinking, Critical Realism and Philosophy: A Confluence of Ideas* (2014), Routledge.

Professor Anne-Wil Harzing

Abstract: “Why Metrics can (and should?) be Applied in the Social Sciences.”

In this presentation I argue that the traditional assumption that metrics cannot be applied in the Social Sciences might in fact have detrimental consequences for the reputation of our discipline. Based on a comparative study of nearly 150 academics in five broad disciplinary areas, I also show that - when using appropriate data sources and metrics - performance of academics across four of the five disciplines is largely comparable. I therefore conclude that as long we use comprehensive databases - such as Google Scholar and Microsoft Academic - and metrics corrected for career length and co-authorship patterns, metrics can and probably should be applied in the Social Sciences.

Bio:

Anne-Wil Harzing (PhD, Bradford) is Professor of International Management at Middlesex University, United Kingdom. Her research interests include international HRM, HQ-subsidiary relationships, the role of language in international business, and the quality and impact of academic research. She has published nearly 100 refereed journals, articles and books/book chapters, and has been listed on Thomson Reuter's Essential Science Indicators top 1% most cited academics in Economics & Business worldwide since 2007. Since 1999 she maintains an extensive website (www.harzing.com) with resources for international management and academic publishing, including the Journal Quality List and Publish or Perish, a software program that retrieves and analyzes academic citations.

Dr Thed van Leeuwen

Abstract: “Metrics and peer review.”

For me, as a researcher in quantitative science studies, and researcher/analyst in numerous studies in applied bibliometrics, accompanying research assessment procedures, the relationship between research metrics and peer review is on the one hand crystal clear, and on the other hand quite complex.

Simple, as we clearly state that research metrics should better not be used in a stand-alone configuration, and this is a position we have taken for decades now. So in the view of CWTS, and we have stated that several times (*Scoping report to the HEFCE*, 2007, and *The Metrics Tide*, 2015), research metrics support and inform peer review, but does not replace peer review. Complex, as it is often difficult to assess to what extent metrics do truly support and inform peers, or that the people judging research performance are not knowledgeable and use research metrics to cover that lack of expertise. Furthermore, peer review and metrics can point into various directions, and the question is how these discrepancies are being dealt with. A next problematic situation relates to the SSH domains, where research metrics do not work, and peer review is the only way of assessing research performance.

Currently, the discussion is on citizen bibliometricianship. A serious risk is here that with respect to research metrics, the attitude is “anything goes!” in this situation, the risk exists that people use research metrics that are not adequate for research assessment purposes. The Internet revolution created a situation in which peer review committee members doing their own inquiries on research metrics, blur their expert perspectives with research metrics of an often dubious quality.

Bio:

Dr. Thed van Leeuwen is a senior researcher at the Leiden University based Center for Science & Technology Studies (CWTS). Originally an Msc in Political Science from the University of Amsterdam, Thed holds a PhD in the Quantitative Studies of Science & Technology from Leiden University. As an academic institute, CWTS serves many clients with their contract research supplying bibliometric data for research assessment procedures.

Thed is appointed at CWTS for over 25 years. As a coordinator and senior researcher, Thed has contributed to over 500 bibliometric studies, which range in level of application from the international level to research group level. As such, the data contained both descriptive bibliometric data as well as more evaluative bibliometric data. Clients Thed has worked for are the European Commission, as well as for Dutch, Belgian, German, Swiss, English, Irish and Spanish national government agencies. Many research institutions used the bibliometric research assessment data produced by CWTS, in the Netherlands the National Research Council (NWO), the Dutch Academy of Sciences (KNAW), and nearly all universities, but also abroad. Furthermore, Thed’s work has contributed periodically to research assessments in the Netherlands under the Standard Evaluation Protocol (SEP) supported by the Dutch Association of Universities (VSNU), NWO and KNAW.

In recent years, his research focused on the assessment of research in the social sciences, the humanities and law, as well as on the issues regarding the Open Science debate. Writer of over 80 journal publications, Thed contributes on a regular basis to the international literature in the field of quantitative science studies, publishing in a variety of journals such as *Scientometrics*, *Journal of the American Society of Information Science & Technology*, *Research Evaluation*, *Research Policy*, *Journal of Information Science*, the *Journal of*

Documentation, *Journal of Informetrics*, *Science & Public Policy*, but also to journals outside of the field of science studies, such as *Nature*, *Cortex*, the *British Journal of Clinical Pharmacology*, *Paediatric and Perinatal Epidemiology,* the *Netherlands Heart Journal, Nederlands Tijdschrift voor Geneeskunde,* and *Medical Teacher*.