

# To intervene, or not to intervene; is that the question? On the role of scientometrics in research evaluation

Opinion piece

Sarah de Rijcke & Alex Rushforth (CWTS, Leiden University)

This is a pre-print of an article accepted for publication in *Journal of the Association for Information Science and Technology* copyright © 2014 (Association for Information Science and Technology)

## Introduction

Recent high-profile statements, criticisms, and boycotts organized against certain quantitative indicators (e.g. the DORA declaration) have brought misuses of performance metrics to the center of attention. A key concern captured in these movements is that the metrics appear to carry authority even where established agents of quality control have explicitly outlined limits to their validity and reliability as measurement tools. This raises a number of challenging questions for those readers of this journal that are implicated in questions of indicator ‘production’ and, by extension, ‘effects’. In this opinion piece we wish to critically engage the question of how producers of indicators can come to terms with their role as (partly-) responsible parties in the current age of ‘evaluative bibliometrics’, without straying towards either uncritical normativism or naïve constructivism (Cronin & Sugimoto, 2014). We do so through the illuminating case of the professional scientometrics community.

We argue that scientometrics is today struggling in its role as one of the ‘regulatory sciences’ (Jasanoff, 1990) as producers of ‘indicators as policy instruments; shaped by regulatory processes’ (Wouters, 1999, 14). In part this struggle is a product of emerging distancing between the increasingly professionalized community and ‘on-the-ground’ research evaluation practices. One of the ways scientometrics can start to strengthen its position and voice within transforming research evaluation contexts is by re-thinking the very foundations of what it considers to be productive ‘intervention’ in these practices.

Quandaries over the strategic positioning of scientometricians in relation to the intervention issue are not new – they have been bubbling for some time. For instance some parts of the community have already been attempting to offer clearer



guidance to end-users and further develop standards for professional use of bibliometrics in research evaluations. Indeed the Science and Technology Indicators (STI) Conference of 2014 has the telling sub-title 'Context Counts'. This title follows on from the 2013 International Society of Scientometrics and Informetrics (ISSI) and STI conferences in Vienna and Berlin, where full plenary sessions were convened on the need for standards in evaluative bibliometrics, and the ethical and policy implications of individual-level bibliometrics (Wouters et al., ISSI newsletter September 2013). The need to debate these issues has come to the forefront in light of reports that certain easy-to-use and potentially misleading metrics for evaluative purposes are becoming a routine part of academic life (e.g. Aksnes & Rip, 2009; Buela-Casal & Zych, 2012; Derrick & Gillespie, 2013; Stephan, 2012; Vanclay, 2011), despite misgivings within the profession itself over their validity (cf. Waltman & Van Eck, 2012; Costas & Bordons, 2007, Moed & van Leeuwen, 1996). At the meetings issues considered included the extent to which this matter now constitutes an ethical responsibility for the community; how it can approach advising on proper uses; and how moving towards clearer, standardized guidelines over usage and consultancy can be achieved. It would be naïve to state that such debates are prompted by altruistic concerns for the satisfaction or welfare of those who use such products, without recognizing also that the community itself has an important stake in mediating and influencing how its products are used. In fact we go as far as to argue that future successes of the field will be told through how effectively this set of strategic problems are negotiated.

We show first how the landscape in which scientometrics has operated under has undergone rapid and unforeseen transformations in the past three decades, becoming much more polyvocal. Inevitably this means more open competition for the attention space of research managers, and as a worst-case scenario the expertise of scientometrics risks being relegated to but one voice in a crowded marketplace. We then wish to problematize how the community has often managed this relationship. Butler (2004) notes conservatism of metric users as a long-standing problem, in displaying a preference for what they consider user-friendly measurements which trump any other inclinations they might otherwise have for adopting more state-of-the art scientometrics. Certainly the growing sophistication and specialization of scientometric techniques, whilst advantageous in many respects, provides challenges with regards actual adoption and diffusion among managers and individuals in research evaluation contexts. But we argue the emerging distance is not simply a product of narrow-mindedness of users. It is also partly due to the role the community ascribes itself in approaching intervention debates. The 'distancing work' frequently performed by scientometricians, we claim, may be counter-productive in the context of intervention strategies. In bringing this particular disjuncture to light, we conclude with suggestions for alternative means for framing how the community might productively re-think the problem of 'intervention', its 'audience', and their own positioning in respect to these challenges.

## The 'implementation problem' in historical perspective

In Europe, Asia, and Australia scientometrics gained prominence in science policy in the 1980s (Garfield, 1979) (Martin & Irvine, 1983) (Narin, 1976). The intertwining of scientometrics and science policy was initially highly effective, although it produced controversies about the relationship that ought to be forged with peer review which remain unresolved to this day (Weingart, 2005). Re-reading the first *Handbook of Quantitative Science and Technology Studies* (1988), it is clear why the nascent profession could be characterized as largely monopolizing the supply of metrics for evaluation purposes. The field had managed to create a demand for their measures - not only by supplying on-demand data and data-handling techniques but also by making sure their products were promoted as policy-relevant information that decision-makers could use strategically (cf. Van Raan & Frankfort 1980). Yet if one compares the optimism in the first handbook with parts of the 2<sup>nd</sup> *Handbook* (2004), certain developments are striking (aside from obvious advances in technical developments and debates). In the intervening period – specifically from around the mid-1990s onwards - disquiet began to be voiced about lack of integration in scientometrics (Glänzel & Schoepflin, 1994), the need for standards (Glänzel, 1996)(Katz, 1996) and the technical efficacy of some metrics developed within scientometrics and widely used for research evaluation purposes (e.g. Garfield, 1996) (Garfield, 2006) (Moed & Van Leeuwen, 1996). Given the field's regulatory ties such debates inevitably spilt-into concerns over adverse 'real-world' effects metrics could have on various fields of scholarship (Seglen 1997, Weingart, 2005). What these voices of discontent signal historically is a growing realization of decreasing power of a field to control and steer uses of its products. By the 2004 Handbook the 'recalcitrance' of policymakers and managers to 'implement' state-of-the-art indicators for evaluations was flagged as a going-concern (Butler, 2004, 403). Since then this 'implementation problem' has not gone away and if anything has extended.

Over the past three decades, the institutional landscape of scientometrics has seen considerable transformations, some of which have strengthened the professional quest for metrics, some of which have eroded its authority. First of all, the community has grown in size and has become more professionalized through membership organizations, regular meetings, conferences, and increasing specialization of journal contributions - all of which characterize maturing of epistemic fields (c.f. Martin et al., 2012). Secondly, there is an increasingly intense globalized competition for expertise in pursuit of promising science and technology. In order to make decisions about how their resources can best fulfill this goal, institutions require all sorts of information-producing tools. This fuels a need for a range of indicators about performance at various levels of aggregation. Thirdly, increasing calls for researchers in public organizations like universities to 'demonstrate their worth' through commercial (and increasingly 'societal relevance')

contributions create further demands for accountability and audit that can assess the extent to which such requirements are (and can be) met. Again various scientometric tools have been designed and adopted to support and drive through such institutional changes. Another source of increased demand and capacities for performance indicators has come from the emergence of new research technologies and their applications. The rise of the internet and digital database and archiving technologies has had a profound impact on new and existing research problems which can be tackled, and the ease, speed, and sophistication at which such tasks can proceed. Clearly the availability of such tools creates new capacities and demands for information on researcher performance. These developments have seen responsibility shifted more-and-more towards individual researchers and appropriation of such indicators by individuals themselves (note for instance the rise of alt-metrics, Google scholar profiles, SciVal tools and so on). Although debates abound whether such metrics depend on and promote the narcissism of its users (Wouters & Costas 2012), its architects have been very effective at circumventing dependence on the authority of traditional elites – including of course the field of scientometrics – overseeing the carrying-out of individual assessments, in ways which have broadly ‘democratized’ this activity, for better or worse. Although there were always commercial and state actors providing an important input into scientometrics, a rise of new actors (Elsevier, Google) and intensified commercial presence of established actors like Thomson Reuters have turned provision of metrics for evaluation purposes into a crowded marketplace (cf. Pontille & Torney, 2013).

Consequences of some of these developments may be positive, resulting in new opportunities for research contributions and information-use, and may increase what the field would consider as effective use of bibliometrics due to more advanced indicators and increased availability of data sets (including web data). Yet some innovations also risk bypassing the quality control mechanisms of fields like scientometrics and the standards they promote. Another troubling consequence of these complex developments has been the rapid and in some cases pronounced consolidation of metrics as research management tools in ways which would surely meet with disapproval from scientometricians, policymakers, users, and stakeholders in public science. In our own recent and ongoing research<sup>1</sup> we find that certain indicators become reified both as research management and as decision-making tools, as both formal and informal standards against which to measure the worth and utility of research activities. While they do not lose their value within the formal order of worth in science, other measures of scientific quality (e.g. originality, long-term scientific progress, societal relevance) either move to the background or become redefined through their relations to quantitative performance indicators.

---

<sup>1</sup> Our research indicates a growing influence of the use of metric-based performance indicators as evaluation and self-evaluation tools on the content and the organization of research, particularly number of publications and impact factors. Project: “*The Impact of Indicators. How Evaluation Shapes Knowledge Production*” (2012-2015; CWTS, Leiden University; dr. Sarah de Rijcke, dr. Alex Rushforth, prof. Paul Wouters).

## **Productive intervention**

The issue of intervention is likely to form a key battle site in the terrain of scientometrics over the coming years. In light of the picture we have painted, it is clear that non-intervention is not a viable solution for the field to entertain. The developments described above give rise to questions not only about validity of new data and data sources, but also where and how the boundary-lines of scientometrics can and ought to be drawn. How might this challenge be approached constructively?

The inability of scientometrics as a professional community to sanction and police what they consider (in-) appropriate uses of metric indicators is what has clearly come to the fore at the ISSI and STI conferences in Vienna and Berlin in 2013. Calls for the community to intervene in order to ensure the integrity of its products were at the heart of the talks (ISSI proceedings, 2013; STI proceedings, 2013). Many contributions qualified that there is indeed a ‘social’ problem, but subsequently either bracketed it off and moved on to deal with unresolved technical issues; zoomed in on epistemic shortcomings and a lack of consensus formation about what scientometric tools actually measure, or pointed out problems with increasingly widespread (mis-) uses of indicators, but without placing any responsibility on bibliometrics. If ethical issues were addressed at all, they were ascribed as deficits of ‘lay’ users’ knowledge and/or of clear ‘professional’ (bibliometrics community) standards - deficits that should be addressed through both technical and ethical guidelines. The gist of the suggestions was to continue working on generic normative first-principles (guidelines, standards) and on improving the epistemic status of measurements, as increasing the level of sophistication would inevitably trickle-down into research management contexts.

It is questionable whether this assumption is correct. Though ‘users’ do need to trust the viability of measures to some degree, ascertaining their epistemic authority will not always be their primary motivation. Moreover, as we already highlighted, scientometricians are currently operating in a very crowded market place. And lastly, the recourse taken here is towards an upstream solution, framed in terms that reinforce one’s own knowledge producing capabilities and professional position. This kind of boundary drawing between professional scientometrics and that of its intended audiences has contradictory effects, because it reproduces the gap between scientometric expertise and the practice level that the community is trying to bridge.

We are not against any kind of distinction of expertise, but an expert division-of-labor based on a linear model of innovation will not suffice. It seems to us that scientometricians are caught in an unproductive model for negotiating difficulties and dilemmas of being a regulatory science, as well as an academic one. Up until recently it has steered towards an optimistic yet technocratic mode of ‘implementation’ (c.f. Todt & Lujan, 2014). The recent surge of concerns and calls for intervention seems to indicate a reevaluation of the interactive relationship between science, science policy, and indicators for science policy. But we would like to stress

that focusing most efforts on producing general, technically sophisticated standards and comprehensible guidelines is not enough. Improving validity and reliability and standards development are obviously useful endeavors, but if scientometricians draw the line of intervention here they risk (at best) maintaining the status quo in terms of how their products get (mis-) used in practice. Instead scientometricians should do well to: a) avoid making *a priori* distinctions between a 'design' and 'implementation' phase within (technical) standards and user guideline development (this will only widen the gap); b) actively involve 'users' in assessing what the problems are and in formulating potential solutions to these problems; and c) include evidence from mixed-methods research to obtain more 'socially robust' (cf. Nowotny, 2003) standards and guidelines. It would be unwise to strictly prescribe what these new dialogic spaces should look like, as this will be something on which the involved parties would have to forge new kinds of pragmatic arrangements. Of course these arrangements need to be assessed in light of the feasibility of certain scientometric interventions. As a scientific field it is not easy (or necessarily constructive) to generate and impose a top-down standard heuristic for its members to follow when approaching intervention. Instead our call is for a re-imagining and re-energizing of discussions and perspectives about how the community should approach the issue. This is all the more important since scientometricians are key actors in the current 'citation infrastructure' (Wouters, 2014), along with publishers, librarians, policy makers, databases, research managers, consultancies, and other metrics users (including scholars). The responsibility for certain applications of bibliometric indicators - from ritualized uses of the JIF to the reification of sophisticated expert metrics - is distributed over all these actors in messy and uneven ways. Scientometricians can potentially draw on a large technical and social-scientific knowledge base (cf. Cronin, 2008), making them well placed to help analyze *and change* the ranges of conceivable types of actions and norms in current practices of research evaluation.

The scientometric community should know by now that simply producing high quality indicators does not ensure seamless passage of its epistemic products into sites of practices to displace inferior products. We contend that the role of 'objective outsiders' who simply produce measures but take no part in their intervention is no longer credible as a mode of intervention. In that particular normative arrangement the field situates itself almost entirely outside the practices of which scientometricians are inherently part. The line between representing and intervening is not so clear-cut, and the history of the scientometric community is a prescient case in point. It has always been very much *in* society, producing indicator knowledge not only for knowledge's sake, but also with a view to intervening in worlds of evaluation policy and practice. The case of scientometrics does not capture exactly the situation of all readers of this journal. But given the increasingly well-documented problems 'users' and 'producers' face in this age of evaluative bibliometrics, it surely brings into sharp relief questions of how professional accountability can and ought to be conceived.



## References

- AKSNES, D. W. & RIP, A. 2009. Researchers' perceptions of citations. *Research Policy*, 38, 895-905.
- BUELA-CASAL, G. & ZYCH, I. 2012. What do the scientists think about the impact factor? *Scientometrics*, 92, 281-292.
- BUTLER, L. 2004. What Happens when Funding is Linked to Publication Counts?, in Henk F Moed, Wolfgang Glänzel, and Ulrich Schmoch (ed.), *Handbook of Quantitative Science and Technology Research*, Kluwer Academic Publishers, The Netherlands, pp. 389-405.
- COSTAS, R. & BORDONS, M. 2007. The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level. *Journal of Informetrics*, 1, 193-203.
- CRONIN, B. & SUGIMOTO, C.R. (2014). Concluding Section. In: Cronin, B. & Sugimoto, C.R. (Eds.). *Scholarly Metrics Under the Microscope: From Citation Analysis to Academic Auditing*. ITI/ASIST (in press).
- CRONIN, B. (2008). The sociological turn in information science. *Journal of Information Science*, 34, 4, 465-475.
- DERRICK, G. & GILLESPIE, J. 2013. "A number you just can't get away from": Characteristics of Adoption and the Social Construction of Metric Use by Researchers. *STI 18th International Conference on Science and Technology Indicators*. Berlin.
- GARFIELD, E. (1979). Is citation analysis a legitimate evaluation tool? *Scientometrics*, 1(4), 359-375. doi:10.1007/BF02019306
- GARFIELD, E. (1996). Fortnightly Review: How can impact factors be improved? *BMJ*, 313(7054), 411-413. doi:10.1136/bmj.313.7054.411
- GARFIELD, E. (2006). The history and meaning of the journal impact factor. *JAMA: The Journal of the American Medical Association*, 295(1), 90-93. doi:10.1001/jama.295.1.90
- GLÄNZEL, W. (1996). The need for standards in bibliometric research and technology. *Scientometrics*, 35(2), 167-176. doi:10.1007/BF02018475
- GLÄNZEL, W. & SCHOEPFLIN, U. 1994. Little scientometrics, big scientometrics ... and beyond? *Scientometrics*, 30, 375-384.
- ISSI Proceedings 2013. <http://issi2013.org/proceedings.html>
- JASANOFF, S. 1990. *The fifth branch : science advisors as policymakers*, London ; Cambridge, MA, Harvard University Press.
- KATZ, J. S. (1996). Bibliometric standards: Personal experience and lessons learned. *Scientometrics*, 35(2), 193-197.
- MARTIN, B. R. & IRVINE, J. 1983. Assessing basic research: Some partial indicators of scientific progress in radio astronomy. *Research Policy*, 12, 61-90.
- MARTIN, B. R., NIGHTINGALE, P. & YEGROS-YEGROS, A. 2012. Science and technology studies: Exploring the knowledge base. *Research Policy*, 41, 1182-1204.

- MOED, H. F., GLANZEL, W. & SCHMOCH, U. 2004. *Handbook of quantitative science and technology research : the use of publication and patent statistics in studies of S&T systems*, Dordrecht ; London, Kluwer Academic Publishers.
- MOED, H. F. & LEEUWEN, T. N. V. 1995. Improving the Accuracy of Institute for Scientific Information's Journal Impact Factors. *Journal of the American Society for Information Science*, 46, 461.
- MOED, H. F. & VAN LEEUWEN, T. N. 1996. Impact factors can mislead. *Nature*, 381, 186.
- NARIN, F. (1976). *Evaluative Bibliometrics: The Use of Publication and Citation Analysis in the Evaluation of Scientific Activity*. National Science Foundation.
- NOWOTNY, H. (2003). Democratising expertise and socially robust knowledge. *Science and Public Policy*, 30, 3, 151-156.
- PONTILLE, D. & TORNAY, D. 2013. La manufacture de l'évaluation scientifique. Algorithmes, jeux de données et outils bibliométriques. *La Découverte/Réseaux*, 1, 177, 23-61.
- SEGLER, P.O. 1997. Why the impact factor of journals should not be used for evaluating research. *BMJ Clinical Research*, 314, 7079, 498-502.
- STI proceedings 2013.  
[http://www.forschungsinform.de/STI2013/download/STI\\_2013\\_Proceedings.pdf](http://www.forschungsinform.de/STI2013/download/STI_2013_Proceedings.pdf)
- STEPHAN, P. E. 2012. *How economics shapes science*, Cambridge, Mass. ; London, Harvard University Press.
- TODT, O. & LUJÁN, J. L. 2014. Values and Decisions: Cognitive and Noncognitive Values in Knowledge Generation and Decision Making. *Science, Technology & Human Values*.
- VANCLAY, J.K. 2011. Impact factor: outdated artefact or stepping-stone to journal certification? *Scientometrics*, 92(2), 211–238.
- VAN RAAN, A.F.J. & FRANKFORT, J.G. 1980. An Approach to University Science Policy: A New Research-Funding System. *International Journal of Institutional Management in Higher Education*, 4, 2, 155-163.
- VAN RAAN, A. F. J. 1988. *Handbook of quantitative studies of science and technology*, North-Holland.
- WALTMAN, L. & VAN ECK, N. J. 2012. The inconsistency of the h-index. *Journal of the American Society for Information Science and Technology*, 63, 406-415.
- WEINGART, P. 2005. Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics*, 62, 117-131.
- WOUTERS, P. (2014). The Citation from Culture to Infrastructure. In: CRONIN, B. & SUGIMOTO, C. (eds.) *Next Generation Metrics: Harnessing Multidimensional Indicators Of Scholarly Performance*. Cambridge, MA: MIT Press.
- WOUTERS, P., GLANZEL, W., GLASER, J. & RAFOLS, I. 2013. The dilemmas of performance indicators of individual researchers – an urgent debate in bibliometrics. *ISSI Newsletter*, September.
- WOUTERS, P. & COSTAS, R. 2012. *Users, narcissism and control: tracking the impact of scholarly publications in the 21st century*, SURFfoundation.





WOUTERS, P. 1999. *The Citation Culture*, Amsterdam, UvA.