

# Bibliometrics: issues and context

Lars Engwall\*<sup>1</sup>, Wim Blockmans†<sup>2</sup> and Denis Weaire‡<sup>3</sup>

\*Department of Business Studies, Uppsala University, Sweden, †Institute for History, Leiden University, The Netherlands, and ‡Institute of Physics, Trinity College Dublin, The University of Dublin, Ireland

## Introduction

Academic institutions have two basic missions: first, to disseminate knowledge (primarily education, but also popularization), and secondly, to create new knowledge (research). Both of these missions include another significant task: quality assessment. In terms of the first mission, academic teachers examine and grade students for their understanding of course material, thereby screening candidates for the labour market [1]. These assessments normally take place inside institutions, although today we can see increasing efforts from authorities to undertake external reviews of standards [2]. Since research is international, such external examinations have long been the case for this second mission. A basic principle for these assessments is the peer-review system, i.e. that the quality of academic work is examined by colleagues in a particular scientific field. This is true for hiring and promotion decisions, assessments of journal manuscripts and research proposals as well as evaluations of institutions.

During the last couple of decades, bibliometrics has come to play a more and more significant role in the peer-review system [3]. This is a result of (i) the development of modern information technology, (ii) a strong expansion of the number of researchers, and (iii) the growth in the number of publication outlets [4]. As a result, bibliometrics has increasingly become the yardstick for quality, a development that has attracted criticism among researchers. In the words of the 1991 Chemistry Nobel Laureate, Richard R. Ernst [5]:

*“We are deeply convinced that human ingenuity and creativity are beyond all conceivable quantitative measure. [...] The present hype of bibliometry made it plainly obvious that judging the quality of science publications and science projects by bibliometric measures alone is inadequate. [...] Start reading papers instead of merely rating them by counting citations!”*

When a Nobel Prize Laureate uses such strong language, the scientific community and science policymakers should pay attention. Although private companies, publishing houses, consultancies and the boards of scientific institutions are organizing perpetuum mobiles of evaluations, the scientific community grows increasingly impatient with the heavy burden imposed by this elaborate evaluation carousel. Whereas the investment in working time and money is evident, the effects are all but transparent and the methodological controversies are far from solved.

<sup>1</sup>Email: lars.engwall@fek.uu.se

<sup>2</sup>Email: wimblockmans7@gmail.com

<sup>3</sup>Email: dweaire@tcd.ie

The main objection is surely the reduction in evaluation to sheer metrics, as a proxy for the desire for a standardized evaluation of a great variety of achievements.

### **Measuring versus quality**

In biosciences and most of the natural sciences, bibliometrics is a well-established practice, widely accepted as a tool in the adjudication of access to scarce resources, on a personal and an institutional level. It is claimed that systemic errors and fallacies such as self-quotation, mutual referencing and negative references may be filtered out, especially in domains with high numbers of researchers. Standardized publication cultures in international journals and a global scientific forum reinforce the applicability of citation indices. They can be used for the ranking of journals, as well as for that of research groups and individual researchers. Collective authorship, however, raises the question of each individual's role, as practices in this respect differ between disciplines.

However, there are several fundamental concerns about the prevailing review systems:

1. To what extent are the current peer-review systems favouring fashionable and standard research, or able to recognize truly groundbreaking ideas? Is there a tension between inherently qualitative characteristics, such as originality and creativity and quantification?
2. Are the current bibliometric systems generally applicable? Engineering, mathematics, computer sciences, as well as most of the social sciences and humanities work with distinctive publication cultures appropriate to their societal mission, forum and target groups. This implies a far greater variety of publication formats in patents, websites, reports, national journals, books and the use of a great number of national languages. In these disciplines, Anglophone researchers may even be uninformed about a considerable body of knowledge published in other languages.
3. Do the self-interested actions of ambitious authors, profit-seeking publishing houses, specialized research bureaus and science administrators lead to systemic distortions of bibliometric measures?
4. Given the amount of the publications under review, and the scope of evaluation committees on an institutional level, are reviewers materially able to assess the quality of research?
5. Is the evaluation work hampering scientific progress by taking time away from research? It is an often-heard argument that in the life of scholars today too much time is taken up by evaluating others and being evaluated by others.

### **Bibliometrics in context**

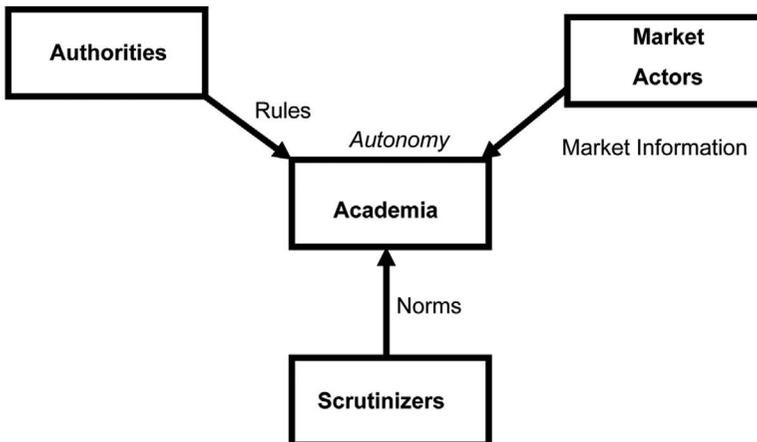
It is appropriate to put bibliometrics into a wider context. In so doing, we can use a more general model of the governance of academic institutions [6]. It implies, as

shown in Figure 1, that organizations are under governance from three types of actors: authorities, market actors and scrutinizers.

Academic institutions have traditionally been characterized by autonomy, i.e. self-governance. There are, in principle, no barriers to starting up academic institutions, but in order to acquire trust they need recognition from authorities (Figure 1, left-hand side). For the medieval universities, this was achieved through bulls from the Pope, and later on, sovereigns provided their recognition [7–9]. This in turn implied certain external rules for the academic institutions, and with the passage of time these have become more forceful as governments have increasingly provided their financial means. More and more, they have been subject to market forces (Figure 1, right-hand side) through various kinds of market information (careers of alumni, publication achievements, successes in grant applications, etc.). These market forces have increased as governments have outsourced their authority to the market to an increasing extent by moving from university block grants to financing through project grants [2]. At the same time, we can also see an increasing role for scrutinizers, i.e. the media, NGOs, auditing bodies and so on (Figure 1, bottom). They provide norms for how the organizations should behave. So, even if academic institutions are characterized by a relatively high degree of autonomy, they are influenced by rules, market information and norms [10].

Although there are a lot of complaints in academia regarding bibliometrics, it was not invented by the authorities. Rather, it is an outcome of an interaction between academia and market actors. There can be no doubt that the present market for publishing is grounded in the desire of scholars to have their results published. This in turn has resulted in the foundation of a large number of journals, often in the form of publications of professional associations. And, these journals have been attractive to publishing houses, since they provide two advantages. First, they are subsidized by the academic community through their willingness to submit and review manuscripts without any remuneration. Secondly, they provide contacts with prospective authors of textbooks, as well as

Figure 1



Academia, authorities, market actors and scrutinizers

with significant gatekeepers in the selection of course literature. No wonder the number of journals is increasing through this symbiosis between academia and market actors! Amidst the resulting plethora of publications and the journals that feed on them, how is one to distil some estimates of quality? Citation analysis, originally designed to assist the research process itself, has stepped into this role.

At the same time it should be noted that, although there are a number of complaints regarding bibliometrics (cf. above), academia itself is not an innocent victim thereof. Instead, citation figures and impact factors are often used by institutions as well as individuals in order to enhance reputations and in order to get access to more resources. So it has even become more and more common to use bibliometrics in the assessment of scholars instead of reading the publications they offer for evaluation. This means an outsourcing to other actors (editors and reviewers) of one of the most important tasks of academia.

The grip of bibliometrics on academia has been reinforced by reactions of the two other types of actors in Figure 1: scrutinizers and authorities. For the scrutinizers, ranking systems constitute a particularly important feature, and such systems pay considerable attention to bibliometrics [11]. Similarly, over the course of time, authorities have shown an increasing interest in bibliometrics as a quality indicator. So, bibliometrics has become a significant part of the governance system outlined in Figure 1.

## **Bibliometrics under scrutiny**

Out of the above critique the following questions arise:

- Do researchers today face an overload of evaluation activities of all kinds on both sides of the process, including journal article and other manuscript reviews?
- Does this time pressure tend to jeopardize the quality of assessments?
- How serious is the bias favouring the English-speaking world?
- Are the humanities and the social sciences, which do not fit well in to the system, seriously mistreated?
- How are we to balance the costs of the evaluation system, the advancement of science and its attendant economic benefits, and the profits made by private companies?

Guided by this critique, the Academia Europaea and the Wenner-Gren Foundations found it appropriate to jointly organize a symposium on the culture of accountability, and the techniques to establish indicators of quality on various levels, from the individual research paper to whole universities. It took place on 23–25 May 2013 in Stockholm, Sweden, and attracted some 60 participants, among them a dozen speakers presenting papers on various aspects of bibliometrics. The present volume contains the revised versions of these papers.

The volume consists of five parts and a conclusion. The first, in addition to this introductory chapter, includes a contribution by Giuseppe Longo, École Normale Supérieure, who delves deep into the philosophy of science, which he

contrasts with democratic mechanisms. He stresses the significance of minority thinking for both democracy and scientific development, which is at odds with ‘democratic’ and ‘normalizing’ thinking as reflected by bibliometrics. He thereby points to the need for editors to find open-minded referees in order to pave the way for original ideas. The next four parts deal with:

- Instruments of measurement
- Indicators for rankings
- Journals, editors and publishers
- Bibliometrics in the humanities and the social sciences

The second part contains two contributions on the techniques for measuring citations and impact. In the first one, Ton (A.F.J.) van Raan, the founding father of CWTS (Centre for Science and Technology Studies) at Leiden University, provides an overview of the application of bibliometrics at different levels of aggregation and in different contexts. He presents a new ‘crown indicator’ and acknowledges 11 issues on which statistical methods should still be refined to avoid pitfalls and sources of error. In addition, he discusses the creation and use of science maps based on word or citation similarity of publications. In an appendix, he provides an overview of the current bibliometric indicators. Further perspectives are provided by Jane Grimson, Trinity College Dublin, who draws parallels between bibliometrics and indicators used in the field of healthcare. In so doing, she discusses the lessons to be learnt from the healthcare sector in assessing research quality. In addition to the current arguments, she demonstrates that “existing approaches to the measurement of research quality are largely gender blind”.

In the third part, the volume turns to methodological critics on the current use of indicators, especially for rankings of universities on a global scale. Michel Gevers, Université catholique de Louvain, presents an analysis of scientific impact versus investments. After a critical appraisal of scientific performance indicators, he compares the performances of a range of countries in an analysis where citation numbers are normalized with respect to the budget invested by that country in scientific research at higher-education institutions. Giovanni Abramo (National Research Council of Italy) and Ciriaco Andrea D’Angelo (University of Rome ‘Tor Vergata’) choose the productivity of the investments in research as their focus. They make a critical assessment of the widely used h-index (Hirsch-index) by applying this measure to evaluate the performance of Italian universities. The authors insist on the necessity to normalize performance indicators by the productivity of individuals, groups or institutions, as well as by the characteristics of any field’s publication styles. Adding up or averaging different datasets leads to distortions and are scientifically incorrect. This analysis leads Abramo and D’Angelo to the conclusion that research evaluations based on the h-index are of little value or may even be dangerous for decision-makers. They are followed by Linda Wedlin of Uppsala University, who discusses how global comparisons matter. She points out how rankings have created an image of a global higher education field, which has not always had positive effects. A counterforce, according to Linda Wedlin, is that the picture is becoming diverse as a result of a multitude of rankings. At the same time she concludes that changes in the system are slow.

Part four discusses the problems of publishers of journals, and their editors. Nicola Gulley, of IOP Publishing, stresses that all parties involved in the process are accountable for results, but she warns against misinterpretation of any single metric taken out of context. Techniques are quickly evolving, such as the download counting of articles, which can be manipulated in the social media. The central question is whether popularity indicates quality. In a subsequent chapter, Jan Reedijk, Leiden University, discusses the value and accuracy of key figures. He equally points to risks of using single parameters due to their own inaccuracies as well as their exposure to manipulation and fabrication by editors and other significant actors. Finally, in part four, Lars Engwall, Uppsala University, elaborates on the quality of quality assessment, with empirical examples pointing to the risks of rejecting important papers as well as accepting fraudulent ones. He also calls attention to increasing difficulties for editors to get colleagues to accept review work.

The fifth part focuses on the particular problems with bibliometric exercises in the humanities and the social sciences. Milena Žic Fuchs from the University of Zagreb deals with the specific problems of bibliometrics in the humanities. She particularly points to the disregard of books, edited volumes and articles in national languages that are common in the humanities. The chapter also presents efforts to handle these problems. In a subsequent chapter, Frederik Verleysen, Pol Ghesquière and Tim Engels, from the University of Antwerp and the University of Leuven, present a database of accredited journals and publishers that was commissioned by the Flemish government in order to include the humanities and the social sciences in benchmarking. They describe the objectives and the design of the system as well as its current content and problems associated with it. They draw a revealing comparison with similar enterprises in three Nordic countries. In the final chapter, Stéphanie Chatelain-Ponroy (Conservatoire National des Arts et Métiers), Stéphanie Mignot-Gérard (Université Paris-Est Créteil), Christine Musselin (Sciences Po) and Samuel Sponem (HEC Montreal) present the results from studies of indicators in French universities, thereby making comparisons between institutions oriented towards the humanities and the social sciences, on one hand, and natural science-oriented institutions, on the other. They find the use and acceptance of the indicators to be higher among natural scientists than in the humanities and the social sciences.

Finally, the editors present some conclusions from the earlier chapters, in the light of both the origins of bibliometrics and of the most recent discussions. We sum them up with ten recommendations. With that we wish readers an enjoyable and stimulating reading of the following chapters. In case you find it appropriate to cite the volume and the contributions therein, we would of course not mind...

## References

1. Blaug, M. (1976) Human capital theory: a slightly jaundiced survey. *Journal of Economic Literature* 14, 827–855
2. Whitley, R., Gläser, J. and Engwall, L. (2010) *Reconfiguring Knowledge Production: Changing Authority Relationships in the Sciences and their Consequences for Intellectual Innovation*. Oxford University Press, Oxford

3. Shatz, D. (2004) *Peer Review: A Critical Inquiry*. Rowman & Littlefield, Lanham, MD
4. Drori, G.S., Meyer, J.W., Ramirez, F.O. and Schofer, E. (2003) *Science in the Modern World Polity: Institutionalization and Globalization*. Stanford University Press, CA
5. Ernst, R.R. (2010) The follies of citation indices and academic ranking lists. A brief commentary to 'Bibliometrics as weapons of mass citation'. *Chimia* 64, 90
6. Engwall, L. (2014) Corporate governance and communication. In *Organizations and the Media. Organizing in a Mediatized World* (Pallas, J., Strannegård, L. and Jonsson, S., eds), pp. 220–233, Routledge, London
7. Lindroth, S. (1976) *A History of Uppsala University 1477–1977*, pp. 12–15, Uppsala University, Uppsala
8. Ruegg, W. (1992) Themes. In *A History of the University in Europe Volume I* (de Ridder-Symoens, H., ed.), pp. 14–20, Cambridge University Press, Cambridge
9. Hammerstein, N. (1996) Relations with authority. In *A History of the University in Europe Volume II* (de Ridder-Symoens, H., ed.), pp. 113–124, Cambridge University Press, Cambridge
10. DiMaggio, P.J. and Powell, W.W. (1983) The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *American Sociology Review* 48, 147–160
11. Wedlin, L. (2006) *Ranking Business Schools: Forming Fields, Identities, and Boundaries in International Management Education*. Edward Elgar, Cheltenham