

Contribution to the Theme Section 'The use and misuse of bibliometric indices in evaluating scholarly performance'



Benefiting from bibliometry

Jarl Giske*

Department of Biology, University of Bergen, Postboks 7803, 5020 Bergen, Norway

ABSTRACT: Society has learned that there can be big pay-offs from science. Governments, private companies and funding agencies will therefore want to know whether their investments in research are well placed. As long as scientists ask for funding support, we must accept the right of these agencies to ask for proof of results. Within a department, the evaluation of performance may be one of several incentives for improving scientific quality and productivity. However, used alone, performance evaluation can lead to destructive competition and marginalization of potentially valuable staff members. Used in combination with research organisation and leadership, it may motivate many staff members. Results-based financing within a department should be used carefully. It should follow the setting of a clear research agenda, by the lowest-level leaders, and the organisation of the activity into groups. Only in groups with goals can the full potential of the staff in a department be mobilised to do more and better science.

KEY WORDS: Bibliometry · Research group · Results-based financing

Resale or republication not permitted without written consent of the publisher

THE SCIENTIST, SCIENCE AND SOCIETY

Many scientists, particularly at universities, feel uncomfortable because of the gradual increase in measures used by their institution or department heads for quantifying their scientific production; such quantification is suggestive of a belief that scientific progress can be counted, or scientific importance can be measured. Universities are generally thought of as arenas for free intellectual pursuit, and should not bear ever increasing resemblance to factories.

Scientists are generally in agreement that the research budgets available to them are far from adequate. Research leaders—from research-group to institutional level—do their best to persuade their peers, industry partners and potential donors to invest more into research every year. Research budgets world-wide have increased over recent decades because scientists and science policy makers have achieved some success in convincing others, notably their governments, that more science is better for society. Three of these arguments are that (1) investment in research is good for the general education and culture of the nation (Freeman 1995), (2) some of the challenges facing society (e.g. environment,

health, development) are so complex that simple solutions will not do (e.g. Lubchenco 1998) and (3) significant investment is necessary to compete in terms of innovation, and in the case of new industries (Pessoa 2005, Madsen 2007).

CAN WE MEASURE SCIENCE?

The short answer is no. There is no easy way of quantifying which of 2 recently published papers will be more influential in terms of (1) setting the future research agendas, (2) resulting in innovations and (3) the number of citations in the scientific literature. However, a number of scientometric indices have been developed with the aim of doing just that. Unfortunately, despite their qualities, they all have flaws or are skewed in one way or another (Seglen 1992, Adam 2002, Butler 2003, Aksnes & Taxt 2004, van Raan 2005), and may even do more harm than good (Lawrence 2007).

Despite the weaknesses of these measures, it may still be possible to make effective use of them. One measure alone cannot provide an unequivocal ranking of the quality or productivity of a group of scientists. Further-

*Email: jarl.giske@bio.uib.no

more, given that science is important for areas as diverse as culture, innovation, and major societal challenges, any single performance measure—however well-designed it may be—cannot capture the overall ‘importance’ of a scientist’s work. A strong performance measure based on a single criterion may actually skew the focus of the staff in an undesirable way, so that they fulfil the criterion at the cost of scientific relevance (Butler 2003, Steele et al. 2006, Lawrence 2007). This proves that, although the use of a performance measure can alter the behaviour of the researchers, a simple measure may do harm as well as good. One could argue that heads of department should not rely on such weak measures, but rather should personally gain insight of the status of work done in the department from reading papers and discussing plans with the staff. However, we should not forget the ability of professors to (rightfully) question such evaluations of themselves and their colleagues, and the loss of focus on their research agenda arising from such conflicts.

I believe that imperfect—or even skewed—statistics can be helpful. However, a handful of measures combined are needed to avoid some of their harmful ‘side-effects’. In my department, we count the number of peer reviewed publications, we sum their journal impact factors, we count the number of master and doctoral theses defended, and we also quantify the teaching and administrative load of each of the staff members. All these measures are easily available, so it does not require huge administrative resources to obtain them. Taken together, I believe that they measure a variety of qualities that are desirable in the staff of a university department. The average research group in my department is given US\$ 6000 per scientist (from PhD student to professor) per year for travels and consumables, but 2/3 of this is based on performance-evaluation measures. In addition, I can divide up US\$ 50–100 000 to serve as start-ups for a few good ideas. By this mix of statistics and personal communication, I believe that we are able to move department resources towards those who use them well.

FROM PERFORMANCE MEASUREMENTS TO MORE AND BETTER SCIENCE

It has been noted by more than one author that a small fraction of the staff in an institution produces the majority of its research and development, whether productivity is measured in terms of publications or patents (Kyvik 1991, Agrawal 2001). Often the same minority is taking on more than its fair share of teaching, supervision and administrative duties. From the perspective of those who provide the funding of my department, this is an obvious argument for an

unequal internal distribution of resources. Yet—and this is my main point—since research and development is so important to society, we cannot leave it to the geniuses; they are too few, and a contribution from the rest of us is also needed.

We should therefore re-examine the reasoning behind performance evaluations. Governments use them to strengthen good institutions. Universities and other institutions use them to boost good departments. It may seem obvious that departments should use them to boost good individuals who will contribute to better output from the department. These individuals are obvious targets for salary increase and better working conditions. But the biggest potential for increased productivity lies in the less productive majority. If the productivity of the department as a whole is to increase, the majority must participate. This is not likely to happen if only the top achievers, as assessed by the performance measurements, are stimulated. I strongly believe in organising the research in teams around the best scientists, so that these top scientists can share their ideas, networks and skills with their colleagues. Even if there are no top achievers around, organising people into groups can help us take advantage of the best qualities in each of us. I therefore believe in maximising effectiveness through a series of parallel actions.

- **Leadership first.** As all performance measures are partly flawed, it is important that the staff is not directed solely towards maximization of such a statistical metric. The head of department must see to it that all staff members have scientific goals. We are not here to top ranking lists, but to address important research questions.

- **Activate human capital to the full.** The largest benefit is obtained by activating all those who normally are not peak producers. Increased focus on the importance of their work can enhance the output of many, but it is important that this is done by stimulation, not by exerting pressure. It is quite easy to measure some statistics, but that will not by itself change much. The challenge is to build better staff members, i.e. to find out how each individual can perform better. For example, there may be small obstacles of some sort which could be removed at relatively low cost. Unfortunately, there are too many employees who feel they have never been sufficiently acknowledged or taken seriously by their immediate superior. Therefore, the first meeting should not be about output statistics. The department head should resemble more closely a team coach than a referee, and must combine the use of performance indicators with an interest in the working conditions of the staff, a recognition of goals reached and a fostering of general enthusiasm for the challenges ahead. Most people, from laboratory tech-

nicians to the world's leading professors, like to hear that their leaders are aware of their efforts. A newsletter or regular department meetings are good arenas. A newsletter may also focus on the less quantifiable achievements (such as persons involved in building good team spirit, or recent appearances in local news media), which still may be important for the long-term goals of the department. By mentioning one good example, many others respond by thinking 'she is not better than me', and will try to prove this. Open approval and admiration can be a means to stimulate others, but there is also a chance that some may be disillusioned about their own achievements if the top achievers get all the attention. In order to administer this within a department of any size, leadership at a lower level is needed: this can be served by the research group

- **Measure groups.** Everyone has natural fluctuations in performance. This will to some extent be evened out if measurements are group-based. By organising the staff into research groups and by measuring these teams, one can foster cooperation at the expense of negative individual competition. Working in teams will allow people to utilise their different talents, so that the group becomes better than the sum of its members. Modern research usually requires a series of skills not often found within any individual brain. Those who may not master the whole research process can still be valuable members of a team. Effective team-building is a job both for the head of department and the research group leader

- **Measure a spectrum of activities.** Since research is important at many levels, make sure that the group is stimulated to activate all its talents. The focus should be not only on journal articles or patents; credit should be given for other important activities such as supervision, teaching, administrative involvements, and even public outreach and networking, as long as the time spent measuring these activities is low. Each measure has its pros and cons; it is important to use several of them

- **Money is not all that counts.** There are many ways to give credit for a good job other than by economic resources. My experience is that internal profiling of staff members through an internal newsletter that reports on a wide range of efforts is rewarding to those mentioned. Further, it creates a positive competitive environment

- **Stability and dynamics.** Science is expensive, both in terms of equipment and personnel. A research group needs both a predictable long-term economic horizon and possibilities for obtaining short-term benefits. I believe that economic incentives should only play a minor role for the internal funding of a research group, at least in areas of high infrastructure and running costs.

From a university department's point of view, the most important benefits that can be achieved through measurement of—and incentives for—scientific production are (1) showing that the leader sees and cares, (2) activation of the less productive part of the staff and (3) stimulation of activity towards long-term scientific goals of the department. To aid the leader in gaining an overall picture, bibliometrics may be helpful, but they alone are not sufficient. The leader must be aware of challenges as well as results, and for all but the smallest departments, the chair will need research-group leaders to bring foresight and compassion into the process. Activating all staff members to work to their potential requires research to be organised into groups. Bibliometrics cannot (and will not) help us set overall and long-term research goals—such as a reduction in poverty, better health or better environment—but it can in a limited way inspire us do better what we already aim at. It is therefore important to combine bibliometrics with a focus on long-term goals. Only when each member is aware that scientific challenges are the justification for the existence of the research group, can a leader use bibliometrics to stimulate the tempo and mode of the group's work.

Acknowledgements. I thank Elinor Bartle, Tove Giske, Randi Taxt and Vigdis Vandvik for constructive suggestions and discussions.

LITERATURE CITED

- Adam D (2002) Citation analysis: the counting house. *Nature* 415:726–729
- Agrawal AK (2001) University-to-industry knowledge transfer: literature review and unanswered questions. *Int J Manag Rev* 3: 285–302
- Aksnes DW, Taxt RE (2004) Peer reviews and bibliometric indicators. A comparative study at a Norwegian university. *Res Eval* 13:33–41
- Butler L (2003) Explaining Australia's increased share of ISI publications: the effects of a funding formula based on publication counts. *Res Policy* 32:143–155
- Freeman C (1995) The 'National System of Innovation' in historical perspective. *Camb J Econ* 19:5–24
- Kyvik S (1991) Productivity in academia: scientific publishing at Norwegian universities. Universitetsforlaget, Oslo
- Lawrence PA (2007) The mismeasurement of science. *Curr Biol* 17:R583–R585
- Lubchenco J (1998) Entering the century of the environment: a new social contract for science. *Science* 279:491–497
- Madsen JB (2007) Are there diminishing returns to R&D? *Econ Lett* 95:161–166
- Pessoa A (2005) 'Ideas' driven growth: the OECD evidence. *Port Econ* 4:46–67
- Seglen PO (1992) The skewness of science. *J Am Soc Inf Sci Technol* 43:628–638
- Steele C, Butler L, Kingsley D (2006) The publishing imperative: the pervasive influence of publication metrics. *Learned Publishing* 19:277–290
- van Raan AFJ (2005) Fatal attraction: conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics* 62:133–143