



THEME SECTION

Lost in publication: how measurement harms science

Peter A. Lawrence*

Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK, and MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK

ABSTRACT: Measurement of scientific productivity is difficult. The measures used (impact factor of the journal, citations to the paper being measured) are crude. But these measures are now so universally adopted that they determine most things that matter: tenure or unemployment, a postdoctoral grant or none, success or failure. As a result, scientists have been forced to downgrade their primary aim from making discoveries to publishing as many papers as possible — and trying to work them into high impact factor journals. Consequently, scientific behaviour has become distorted and the utility, quality and objectivity of articles has deteriorated. Changes to the way scientists are assessed are urgently needed, and I suggest some here.

KEY WORDS: Impact factors · Citations · *h*-index · Measurement · Scientific careers

—Resale or republication not permitted without written consent of the publisher—

'Are all thy conquests, glories, triumphs, spoils,
Shrunk to this little measure'

William Shakespeare: *Julius Caesar*, Act III, scene I

'What category? Minor writer? ... Significant minority writer? Major minor writer? ... never minor major writer? What did he say to you?'

Brian Friel, 1997: *Give Me Your Answer, Do!* (p. 48)

'Every now and then I receive visits from earnest men and women armed with questionnaires and tape recorders who want to find out what made the Laboratory of Molecular Biology in Cambridge so remarkably creative. They come from the social sciences to...seek their Holy Grail in interdisciplinary organisation. I feel tempted to draw their attention to 15th-century Florence with a population of less than 50,000, from which emerged Leonardo, Michelangelo, Ghiberti...and other great artists. Had my questioners investigated whether the rulers of Florence had created an interdisciplinary organisation of sculptors, architects, and poets to bring to life this flowering of great art?...My questions are not as absurd as they seem, because cre-

ativity in science, as in the arts, cannot be organised. It arises spontaneously from individual talent. Well-run laboratories can foster it, but hierarchical organisation, inflexible, bureaucratic rules, and mounds of futile paperwork can kill it. Discoveries cannot be planned; they pop up, like Puck, in unexpected places.'

Max Perutz: Preface to *I Wish I'd Made You Angry Earlier: Essays on Science, Scientists, and Humanity*

No matter what measures are devised to assess people, no matter how inadequate they are, it is human nature to try to meet them and to aim to excel. It should, therefore, be no surprise that attempts to quantify the quality of scientific publications have produced many behavioural and political changes in the world of science.

It has always been crucial for research scientists to publish their work. There have always been 3 purposes: first, to disseminate new information so that others can learn from it; second, so that other scientists may repeat the studies, or build on them with additional observations or experiments; and only third, so that the support, financial or otherwise, for the scientist can be justified to interested parties. This third reason used to be subsidiary, but no longer; publication has

*Email: pal@mrc-lmb.cam.ac.uk

become the main goal because it is the scientist's life-line (Lawrence 2003). This enormous change in emphasis has damaged the practice of science, has transformed the motivation of researchers, changed the way results are presented and reduced the accuracy and accessibility of the scientific literature.

Let me explain: if you need to publish because you need a meal ticket, you must publish when you are hungry, not when the work is completed or when it is comprehensible or valuable. This explains, for example, why many PhD projects are published in an unsatisfactory way before the end of the PhD training: if the student is to stay in science after their degree is awarded they will need a grant and most grant applications require publications. These publications take time to prepare and to be accepted by journals, and this has to be added to the delay from the initial grant application to when the money comes in—meaning that many students must start preparing papers so soon that they can hardly know what they are doing! Post-docs must publish in order to get a job, and senior scientists must steadily add to their list of papers in order to renew grants or get tenure. Time and again work is submitted prematurely to journals, or projects are urgently cobbled together, or specially constructed to 'make papers' before deadlines. Skilled career scientists are learning how to package up their work in parcels of the right size so that they can be fed first to the journals and then into the maws of the granting agencies.

Then there is the question of which journal to send the work to. Since scientists are now assessed, not so much by the validity, interest or quality of the work itself, but by the impact factor of the journal (Steele et al. 2006), many, if not most scientists, spend too much time and effort thinking and worrying about publication strategy. What should one do if the paper is rejected by the journal of first choice? Is it worth having another go at the editor in the hope of changing their decision, or would it be more prudent to try another journal immediately? If another journal, then should one just send the paper off again unchanged, or should one try to put the results into a different kind of fancy dress? Maybe some other finding can be brought into the paper to give it a more up-to-date buzz. Perhaps the paper needs a link to something trendy published recently, a 'hook' as journalists call it, or a connection to a current controversy: scientists have learnt that editors also have to maintain the impact factor of their journals, and controversies can generate citations. There are so many ploys that can be tried, but they all take time and if they fail, as they often do, months are wasted and priority can be lost.

These stratagems are aimed first at editors, as it is she or he who will be deciding whether the paper

should be reviewed. However, once past that hurdle, the next obstacle is the review process, and reviewers present different dangers for the skilled paper writer. One must tread very carefully here. One must make sure to quote all the 'leaders' of the field. It may well be safer to play down any results that do not fit with the perceptions of likely reviewers and may be unwise to confront contrasting results openly. The normal strategy nowadays is to ignore or hide conflicts, for bringing them into the open may be risky, as perhaps one of the reviewers may have some stake in these results (or even be responsible for them) and might be irritated. This is especially pertinent with the high impact factor journals, as editors of these journals regularly insist that the paper be praised 'enthusiastically' by every one of the reviewers and, in consequence, one equivocal review can sink the paper. However, all of these stratagems have their downside, as they usually reduce the objectivity and clarity of the arguments. Politics enervates science.

In my opinion, the growing use of the *h*-index (Hirsch 2005)—which, based on numbers of citations, tries to quantify both the productivity and impact of a scientist—will lead to a dependence on, and an obsession with, citations. I predict that 'citation-fishing and citation-bartering' will become common practice (Lawrence 2007). Note that most journals are now limiting the number of references, either directly or by restricting the number of words in articles and including references in that count. Authors are thus encouraged to make inadequate or inappropriate citations. For example, it takes fewer words to refer to a single review than to several original papers that report a finding, and, since there is usually a choice of several different relevant reviews, the selection of which review to cite can easily become more political than scientific. However unjust citations may be, scientists will evaluate their own work by the citations they receive and this will lead to dubious practices. For example, scientists will claim superiority over others if they have more citations, and this will be endorsed by bean counters everywhere. As a consequence, those practising in small fields (where papers tend to be cited less) will be discouraged.

David Colquhoun's case study of Imperial College, London, analysed the consequences of embracing these measures (Colquhoun: www.dcs.science.net). In that apparently benighted college each scientist is expected to produce at least 3 papers per year in which at least 1 should be in a journal with an impact factor above 5. Colquhoun persuasively argues that these requirements make no sense, pointing out, for example, that if the current measures had been applied in the past, some of their very best scientists, including Nobel Laureates, would have failed from time to time in their careers.

There are other consequences of the use of numerical measures: given that meeting them rewards aggressive, acquisitive and exploitative behaviour (Lawrence 2002, Lawrence 2003, Montgomerie & Birkhead 2005), their use will select for scientists with these characteristics. I have argued elsewhere that creative discovery is not helped by measures that select for tough fighters and against more reflective modest people (Lawrence 2006). For example, Ed Lewis, winner of the 1995 Nobel Prize in Physiology or Medicine and throughout his life a dedicated and unassuming scientist, as well as a rare and irregular publisher, had a very low *h*-index and might not win grants in today's world. Furthermore, on average, women are less aggressive than men (Baron-Cohen 2003, Babcock & Laschever 2004, Lawrence 2006, Symonds et al. 2006) and thus, although these measures discriminate against gentler scientists of both genders, more women are affected. Statistics leave no doubt that women most frequently drop out of bioscience (ETAN 2000), and many of them are very talented. The most frequently voiced reasons are (1) 'I don't want to spend the rest of my life competing and showing off', (2) 'I don't want to take up prize fighting to get grants', and (3) 'modern science has become a struggle for survival, and I would rather do something else'.

What should be done to improve things? In my opinion, grant agencies ought to use other ways of evaluating the quality of work, and declare that they are doing so. Those they support might then spend less time scheming to get their papers into vanity journals and more time researching and solving scientific problems (Keller 2007). In my limited experience, grant applications do not describe what you will actually do but are in reality an ingenuity and knowledge test in which honesty is little valued; they amount to an attempt to demonstrate that one knows what one is doing and can divine what the outcomes of experiments will be and assess what might be risky to reveal. One needs a shrewd idea of what might impress the assessment committees. I am not sure how well time or science is served by this rather weird process. To reduce the pressure on scientists who spend so much time concocting grant applications and to reduce the number of screeds that have to be assessed and compared, I suggest offering an alternative means of grant application that would be based entirely on the quality of the last 5 years' work as is the case in Canada (CNSRC). All the applicant would have to do, if they chose this option, would be to list say 3 of what they think are their best papers from the last 5 years and describe briefly who was responsible for the work therein. An advantage of

having a small number of papers to assess, apart from the obvious one of less to read, is that authors would be encouraged to write a smaller number of papers that mean and convey something important, rather than a large number of papers to be counted. It is encouraging that in the USA and specifically for the Howard Hughes Medical Institute assessment is limited to 5 papers from the last 5 years, chosen by the applicants. Perhaps this is the first step in what I think would be a useful trend. But how should one evaluate these papers? Just by reading them: there is no better way. This method was used in the past; it is not 'objective', but should be an attempt to get at what matters, better I think than relying on a precise measurement of what does not matter. Research has to be evaluated for rigour, originality and significance, for the light it casts and for economic and heuristic value; these qualities may be difficult to assess, but we should try.

LITERATURE CITED

- Babcock S, Laschever S (2004) *Women don't ask: negotiation and the gender divide*. Princeton University Press, Hoboken, NJ
- Baron-Cohen S (2003) *The essential difference. Men, women and the extreme male brain*. Allen Lane, London
- ETAN (European Technology Assessment Network) (2000) *European Technology Assessment Network on women and science, 'Science policies in the European Union: promoting excellence through mainstreaming gender equality'*. European Commission, Brussels
- Friel B (1997) *Give me your answer, do!* The Gallery Press, Loughcrew
- Hirsch JE (2005) An index to quantify an individual's scientific research output. *Proc Natl Acad Sci USA* 102: 16569–16572
- Keller L (2007) Uncovering the biodiversity of genetic and reproductive systems: time for a more open approach. *Am Nat* 169:1–8
- Lawrence PA (2002) Rank injustice. *Nature* 415:835–836
- Lawrence PA (2003) The politics of publication. *Nature* 422:259–261
- Lawrence PA (2006) Men, women, and ghosts in science. *PLoS Biol* 4:e19, doi:10.1371/journal.pbio0040019
- Lawrence PA (2007) The mismeasurement of science. *Curr Biol* 17:R583–R585
- Montgomerie B, Birkhead T (2005) A beginner's guide to scientific misconduct. *Int Soc Behav Ecol* 17:16–24
- Perutz M (2003) *I wish I'd made you angry earlier: essays on science, scientists, and humanity*. Cold Spring Harbor Laboratory Press, Woodbury, NY
- Steele C, Butler L, Kingsley D (2006) The publishing imperative: the pervasive influence of publication metrics. *Learned Publishing* 19:277–290
- Symonds MRE, Gemmell NJ, Braisher TL, Gorringer KL, Elgar MA (2006) Gender differences in publication output: towards an unbiased metric of research performance. *PLoS One* 1:e127, doi:10.1371/journal.pone.0000127

Editorial responsibility: Howard Browman, Storebø, Norway and Konstantinos Stergiou, Thessaloniki, Greece.

NOTE: Howard Browman's finalization of editorial work on this Theme Section is appreciated. It does not affect the termination of his cooperation with Inter-Research on Nov 5, 2007

*Submitted: August 29, 2007; Accepted: November 10, 2007
Proofs received from author(s): January 14, 2008*