

THEME SECTION



Escape from the impact factor

Philip Campbell*

Nature, 4 Crinan Street, London N1 9XW, UK

ABSTRACT: As Editor-in-Chief of the journal *Nature*, I am concerned by the tendency within academic administrations to focus on a journal's impact factor when judging the worth of scientific contributions by researchers, affecting promotions, recruitment and, in some countries, financial bonuses for each paper. Our own internal research demonstrates how a high journal impact factor can be the skewed result of many citations of a few papers rather than the average level of the majority, reducing its value as an objective measure of an individual paper. Proposed alternative indices have their own drawbacks. Many researchers say that their important work has been published in low-impact journals. Focusing on the citations of individual papers is a more reliable indicator of an individual's impact. A positive development is the increasing ability to track the contributions of individuals by means of author-contribution statements and perhaps, in the future, citability of components of papers rather than the whole. There are attempts to escape the hierarchy of high-impact-factor journals by means of undifferentiated databases of peer-reviewed papers such as PLoS One. It remains to be seen whether that model will help outstanding work to rise to due recognition regardless of editorial selectivity. Although the current system may be effective at measuring merit on national and institutional scales, the most effective and fair analysis of a person's contribution derives from a direct assessment of individual papers, regardless of where they were published.

KEY WORDS: Bibliometrics · Impact factor · Academic merit

Resale or republication not permitted without written consent of the publisher

INTRODUCTION

As a competitive editor, I focus on (among many things) the added value that my colleagues provide to authors in enhancing their papers, and to readers in the assessment of a paper's importance. Citations are one measure of our success, and our papers are indeed highly cited, but I pay as little attention to the 'impact factor' as possible.

For many years, I have been expressing concern about the dependence by many on the impact factor as a measure of esteem. At the risk of rehearsing well understood considerations, I will summarise my worries before being more constructive about other ways forward.

CONCERNS

The principal concern I have is about crudeness. In 2005, we measured this. The value of *Nature's* impact

factor for 2004 was 32.2. As discussed in an editorial (*Nature* 435:1003–1004; 2005), we analysed the citations of individual *Nature* papers over the relevant period (i.e. citations in 2004 of papers published in 2002 to 2003) and found that 89% of our impact factor was generated by just 25% of our papers.

The *Nature* paper most cited in 2004, among those published during 2002 to 2003, was the mouse genome. That paper represented the culmination of a great cumulative enterprise, and as such was an important point of reference rather than an expression of unusually deep mechanistic insight. When we analyzed the numbers in 2005, it had received >1000 citations. Within the impact-factor measurement year of 2004 alone, it received 522 citations. Our next most cited paper from 2002 to 2003 concerned the functional organization of the yeast proteome and received 351 citations in 2004. Only 50 out of the roughly 1800 citable items published in 2002 to 2003 received >100 citations in 2004. The majority of our papers received fewer than 20 citations.

*Email: nature@nature.com

The impact factor is taken by some administrators as a measure of the typical citation rate for the journal, but as these values suggest, for many journals, it isn't 'typical' at all.

Why were most papers' citations 'low' (by our standards)? Because they were in disciplines with characteristically low citation rates per paper like physics, or with citation rates that are typically slow to grow, like the Earth sciences, or because they were excellent (e.g. visionary) but not 'hot'.

The impact factor has another aspect that I heartily dislike. In 2003 and 2004 we published significantly more papers than our principal competitor, *Science*. But the impact factor's formula divides the number of ISI-counted citations of 2 years' worth of *Nature* papers by the number of *Nature*'s 'citable items' in that same period. The larger the number of papers, the lower the impact factor. In other words, worrying about maximizing the impact factor turns what many might consider a benefit — i.e. more good papers to read — into a burden.

To add injury to insult, the numbers quoted in calculating the impact factor are highly questionable. Try as we might, my colleagues and I cannot reconcile our own counts of citable items in *Nature*, several other *Nature* journals and indeed *Science*, with those used by ISI.

Some colleagues care a lot about this competitive ranking, even though they fully realise how arbitrary the number is. It does, after all, have an impact on researchers' perceptions of journals. I cannot deny that our own marketing department celebrates the high rankings of *Nature* journals but, for what it's worth, I care about it only to the extent that anybody might be influenced by these numbers in a way that adversely affects us. The editors of *Science* have every right to be proud of their papers, and of the spectrum of citation levels they achieve. I happen to think we tend to publish better papers, but that word 'better' begs questions and subjectivities. To what extent we are succeeding is for readers to decide on their own terms, but the judgement of 'better' is best kept independent of the impact factor.

Some may see this as disingenuous posturing. After all, they might say, *Nature* sits alongside other journals and the impact factor itself as keystones of a tyranny of academic accountability, in which everyone feels compelled to jostle for publication in high-impact journals. I say that the journals are doing their job in selecting strong papers, and that how others exploit this activity to their own ends is up to them.

It is possible to be more analytical. For example, the biologist Peter Lawrence has pointed to behavioural aspects of this 'tyranny', including a growth in inappropriate honorary authorship, increased authorships of papers to an extent that hides true contributions, and

otherwise-unnecessary travel to meetings to schmooze with (among others) editors of journals (Lawrence 2003, 2007). My own view is that anyone who believes that this last activity will significantly enhance their likelihood of publication is deluding themselves.

Any measurement system will produce gamesmanship to increase success. Lawrence yearns for the one system that obviates it: face-to-face conversations between peers that allows a truly deep and well focused estimate of scientific merit. I do too. It's a shame that this is so often considered impractical.

POSSIBLE ALTERNATIVES

Recognising the necessity for quantification as an input in assessing impact, many people are investigating better measures, as can be seen in the pages of *Scientometrics* and other journals. The Wikipedia article on Hirsch's h-index highlights its shortcomings, yet many now deploy it. An article comparing such indices highlighted the virtues of counting the mean number of citations per paper in assessing an individual scientist's impact (Lehman et al. 2006).

I am intrigued by the possibility of greater granularity within the literature. This is already happening in relation to authors' contributions. In *Nature* and *Nature* research journals (following the lead taken by medical journals, to their great credit), we encourage authors to include brief summaries of which author contributed what to the work. More and more author collaborations are taking up this option. This will be spurred on and will no doubt become more formal if funding agencies and institutions begin to explicitly track such information.

Another change that is possible though yet to emerge is to enable the citation of sub-sections of papers. This requires these elements (or 'articles' particles' as a colleague calls them) to be individually identifiable, e.g. with a Digital Object Identifier (DOI). Recently 2 researchers — one an ecologist and one a microbiologist — mentioned to me how their respective highest cited papers were in obscure journals, and that the reason for the citations was not so much the main conclusion of the paper but rather a technical aspect of the methods used. It would surely be healthy to allow people the option of citing these sub-elements of papers, especially in an envisaged machine-enhanced literature in which computers have extracted such key elements in sophisticated search and aggregation. The same should be possible for contributions in the grey literature, such as individually contributed segments of databases that turn out to be of great value.

Such enhancements to the system will not do away with one key aspect of the accountability tyranny: the

over-dependence, in superficial assessments, on journal brands, and especially the excessive emphasis on those with the very highest impact factors. One way forward, again increasing the granularity of available information, is to make citation statistics for every paper more readily available.

There is an experiment under way that attempts to do away with prestige brands in the literature altogether; a refereed archive that makes no attempt to select papers for excellence. Of such a kind is 'PLOS One', the online journal published by the Public Library of Science (www.plosone.org).

The aspect that most excites its founders is the ability to comment on its published papers. As one expects from these publishers, it is based on the author-pays model. Both of these features are experiments in their own right, and have yet to prove themselves, respectively, in popularity or financial viability. What is more to the point here is that the PLoS One model deliberately avoids the very mission by which *Nature* tries to compel the attention of its readers every week—selecting the most interesting science from whatever is sent to it. Instead, the only criterion is that a paper has been judged by referees to be scientifically credible.

However, nobody wants to have to wade through a morass of papers of hugely mixed quality, so how will the more interesting papers in such an archive get

noticed as such? Perhaps the well-established preprint server arXiv (<http://arxiv.org>) provides a model. It contains thousands of pre-publication draft papers and makes no claim about their significance, but the only way people can spot content of particular interest is to visit it every day and see what's been added to the relevant subject area.

Readers won't be surprised if I stand by the model in which editors select interesting papers, for all its shortcomings. I have indicated here some ways in which the use of associated statistics might be improved.

I am sure that citation statistics of large numbers of individual papers can reflect the impact of contributions at the institutional or national level. They can also provide useful input at the level of an individual. But for a sure assessment of an individual, there is truly no substitute for reading the papers themselves, regardless of the journal in which they appear.

LITERATURE CITED

- Lawrence P (2003) The politics of publication. *Nature* 422:259–261
Lawrence PA (2007) The mismeasurement of science. *Curr Biol* 17:R583–R585
Lehmann S, Jackson AD, Lautrup BE (2006) Measures for measures. *Nature* 444:1003–1004

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

*Submitted: September 14, 2007; Accepted: November 10, 2007
Proofs received from author(s): December 17, 2007*