CURRENT SCIENCE

Volume 96 Number 10 25 May 2009

EDITORIAL

Metrics of Science: Loosening the Stranglehold

Counting citations, flaunting *h* indices and advertising average impact factors of publication lists have become a common practice amongst members of India's growing community of scientists and researchers. Modesty and reticence do not seem to be desirable qualities for career advancement in science, worldwide. The reclusive laboratory scientist is an extinct species, driven into oblivion by the rise of the aggressive, managerial professor. The administration of very large research teams and the need for constant projection of even preliminary results in conferences across the globe requires managerial and marketing skills, more suited to the world of commerce. In a recent lecture I heard a distinguished colleague looking back on research done decades ago, proclaiming wistfully at the end: 'it was fun'. Even as we collectively lament on the lack of enthusiasm of young students for science, there is little discussion of how our institutions and the researchers within them are perceived from the outside. 'Fun' may be a word that is hard to associate with a scientific community, obsessed with quantitative performance parameters; reminding cricket lovers of batsmen who play with an eye on averages. My thoughts turned to citation analysis following a recent episode where the Indian Academy of Sciences asked for scientometric parameters to be provided with the nomination form of candidates being considered for election to its Fellowship. Shortly after, the requirement was withdrawn, presumably because it seemed unbecoming of a body that should be able to assess scholarship by studying all the supporting material provided by the nominators. I would not have given a second thought to this episode if a colleague, distinguished by election to the Academy, had not walked into my office and charged me with some responsibility for the Academy's change of stance. He said somewhat accusingly: 'You have written favourably about the metrics of science, on occasion, and at other times dismissively about the obsessive use of citation counts. It is such fickle opinion that results in confusing readers'. I was momentarily flattered that someone should even think that casual essays can mould opinion. However, my reaction was somewhat cautious and I advanced the defence that none of my columns had promoted the use of citation statistics as a robust measure for assessing individuals. To further buttress my case, I argued that in any event famous moulders of public opinion had noted that 'consistency is the virtue of an ass'. But my critic was in no mood to listen, walking out scarcely listening to my mild expression of protest. My voluble critic seemed to like the *h*-index and seemed annoyed at the Academy's reluctance to endorse its use for evaluating scientists. Even as I wondered about the epidemic proportion of the obsession with scientometric indicators within the Indian scientific community, my attention was drawn to a new parameter with an impressive name – the eigenfactor. Having waded, somewhat unsuccessfully, through a world of eigenvalues and eigenvectors in my student days, this seemed a term that would undoubtedly be a force to reckon with; sweeping through the scientific community with a degree of infectious success, that would shame even the most prolific virus.

Over the years as a bemused observer of both scientists and the expanding discipline of scientometrics, I have often wondered about their relationship. Professional practitioners of scientometrics seem dedicated observers of science, developing parameters and methods to quantitate scientific output and impact. Scientists, their journals and their institutions use these parameters largely for self promotion, rarely for introspective analysis. Most professional scientometric analysts have little feeling for science itself, hoping that insights might emerge from impersonal quantitative methods. In recent times practising scientists, often physicists, have entered the field of scientometrics; bringing with them increasingly complex ways of analysing the exploding volume of scientific literature. In returning to the topic of the metrics of science, I must confess that over the last eight years or so this is a topic which has engaged me on half a dozen occasions. Most recently, a column was entitled 'Scientometrics: A Dismal Science' (*Curr. Sci.*, 2008, **95**, 431–432), clearly positioning me on the side of those who view quantitative measures with some suspicion. Despite my prejudices, I was drawn to a commentary irresistibly entitled 'The most influential journals: impact factor and eigenfactor' (Fersht, A., *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 6883). The author asks if 'there is a simple statistical metric of the influence of a journal?'. He goes on to dismiss the 'impact factor' and 'total citations' arguing that both measures are affected by the total number of papers published in the journal. He then promotes the 'eigenfactor', a ranking procedure that is 'similar to that used by Google for ranking the importance of Web sites in a search'. Readers who wish to jump on the eigenfactor

bandwagon might like to read papers emanating from a biology department, which explain the new metric (Bergstrom, C. *et al.*, *J. Neurosci.*, 2008, **28**, 11433; Bergstrom, C. and West, J. D., *Neurology*, 2008, **71**, 1850). There is a strong correlation between the total number of citations and the eigenfactor, with *Nature*, *PNAS* and *Science* emerging as the journals with 'the most overall influence on science'. Biochemists and molecular biologists with a traditional upbringing will be pleased to note that the *Journal of Biological Chemistry* (*JBC*), an old warhorse, appears higher ranked than *Cell*, reversing their perceived importance. Unsurprisingly, *PNAS* chose to publish this analysis; an indication that even staid and venerable journals need to self-consciously promote themselves in a fiercely competitive world of publishing. If the 'eigenfactor' is a good single number indicator of a journal's influence, can the 'average eigenfactor' of a list of publications be far behind? Fersht worries about 'the terrible legacy' of the impact factor that is 'being used to evaluate scientists rather than journals'. He concludes that 'the least evil of the metrics of individual scientists is the *h*-index'.

When did the practice of ranking journals originate? Clearly, this must be a subject of some importance for librarians with limited budgets. An early attempt is described in a paper entitled 'College Libraries and Chemical Education', which addressed the issue of providing journals 'necessary for the stimulation and intellectual development of the faculty' in an undergraduate college. The authors noted that journals were needed because of 'the demand of the colleges for instructors with the doctorate degree. Such men are reluctant to accept positions in colleges where facilities for continuing the research which they have learned to love are lacking' (Gross, P. L. K. and Gross, E. M., *Science*, 1927, **66**, 385). This report, which was limited to chemistry journals, used a simple procedure. They selected *The Journal of the American Chemical Society* (*JACS*) as a standard ('most representative of American Chemistry') and counted citations in *JACS* to other journals as a measure of influence. The ranking that appeared over eight decades ago must interest analysts, with a weakness for history. The top five included three German journals (*Berichte*, *Annalen* and *Zeit. Physik. Chem.*), one French journal (*Compt. Rend.*) and only one English journal (*J. Chem. Soc.*). Many of these journals have disappeared, as the face of science has changed dramatically over the intervening years. Interestingly these authors discuss criteria for the purchase of 'back files', noting that the *JBC* should be provided for 'students looking forward to the study of medicine'. The authors noted in 1927 that such an analysis would be helpful in other fields. The subsequent introduction of the journal impact factor revolutionized the field of ranking journals. Ironically, libraries and librarians appear to be minimally interested in impact factors. Instead it is scientists who now use impact factors as a means of judging other scientists. As criticism of journal impact factors as a criterion of judgement has mounted, new performance indices based on citation counts have appeared, with the *h*-index currently ruling the roost. The area of scientometrics is now in the grip of two powerful commercial interests – Thomson Scientific, which now markets the *Web of Science* and Elsevier, which promotes the *Scopus* database. Clever marketing strategies, including awards for 'citation laureates' in chosen countries have increased the number of scientists obsessed and addicted to citation counting. It is common to see Ph D students and postdoctoral fellows choosing journals to send their papers, based on impact factors. Gone are the days when one 'just knew' the most desirable journals in a field. Since the highest ranked journals are often unattainable, impact factors are used to make choices between what are considered 'mid-level' journals. In many ways this is an undesirable practice; personal judgements may at least permit students to come to their own conclusions.

The field of measuring scientific impact continues to grow. A formidably titled analysis, 'A principal component analysis of 39 scientific impact measures' begins engagingly: 'Science is a gift based economy; value is defined as the degree to which one's ideas have contributed to knowledge and impacted the thinking of others'. The conclusion of this study should be music to those who worry about the impact of the impact factor: 'The commonly used citation Impact Factor is not positioned at the core of this construct, but at its periphery, and should thus be used with caution' (Bollen, J. *et al.*, arXiv: 0902.2183vl, 12 February 2009). The age of the Internet has also led to the use of the number of 'downloads' to rank articles (and their authors), although it seems apparent that downloading, reading, digesting and citing are distinct activities that may reflect the influence of a paper on a reader. Despite many warnings over the years, scientists and managers of science in India increasingly use measures like impact factors and *h*-indices to make comparisons of individuals, journals and institutions with little regard for differences between disciplines. A recent analysis examines differences across fields and draws attention to 'inflation' in impact factors over time (Althouse, B. M. *et al.*, *J. Am. Soc. Inf. Sci. Tech.*, 2009, **60**, 27). The average impact factors for journals in defined fields vary widely. The numbers for some representative fields are: Molecular and Cell Biology 4.76, Chemistry 2.61, Physics 1.91, Chemical Engineering 1.29, Fluid Mechanics 0.80, Materials Engineering 0.82, Mathematics 0.55). Clearly, cross-discipline comparisons must be avoided and inter-institutional comparisons may turn out to be flawed.

Evaluating science is probably best done, even today, by reading and listening. Authors must write and readers must read, without self-consciously worrying about the status of journals, to the point where it becomes a distracting obsession. Administrators cannot abdicate a responsibility for judgement, when selections and promotions are involved. If the stranglehold of the *h*-index and the impact factor is loosened science might again be fun.

P. Balaram