



THEME SECTION

The siege of science

Michael Taylor^{1,*}, Pandelis Perakakis², Varvara Trachana³

¹Departamento de Astrofísica Molecular e Infrarroja (DAMIR), Instituto de Estructura de la Materia (IEM), Consejo Superior de Investigaciones Científicas (CSIC), Calle Serrano 121, Madrid 28006, Spain

²Departamento de Psicología, Campus Universitario de Cartuja, Granada 18071, Spain

³Departamento de Inmunología y Oncología, Centro Nacional de Biotecnología (CNB), Darwin 3, Campus de Cantoblanco, Madrid 28049, Spain

ABSTRACT: Science is in a state of siege. The traditional stage for scientific ideas through peer-reviewed academic journals has been hijacked by an overpriced journal monopoly. After a wave of mergers and take-overs, big business publishing houses now exercise economic control over access to knowledge and free scientific discourse. Their 'all is number' rationale, made possible and perpetuated by single-parameter bibliometric indices like the Impact Factor and the h-index has led to a measurement of scientists, science and science communication with quality being reduced to quantity and with careers hanging in the balance of column totals. Other multi-parameter indices like the subscription-based Index Copernicus have not helped to resolve the situation. The patented and undisclosed black box algorithm of the Index Copernicus has just replaced one yardstick by another even less accessible one. Moreover, the academic as author, editor and/or reviewer, under intense competitive pressure, is forced to play the publishing game where such numbers rule, leading to frequent abuses of power. However, there are also deep paradoxes at the heart of this siege. Electronic software for producing camera-ready-copy, LaTeX style files, the internet and technology mean that it has never been easier or cheaper to publish than it is today. Despite this, top journals are charging exorbitant prices for authors to publish and for readers to access their articles. Academic libraries are feeling the pinch the most and are being forced to cut journal subscriptions. Not surprisingly, scholars in droves are declaring their independence from commercial publishers and are moving to open access journals or are self-archiving their articles in public domain pre-print servers. That this movement is starting to hurt the big publishing houses is evidenced by their use of counter-tactics such as proprietary pre-print servers and pure propaganda in their attempts to guard against profit loss. Whether or not bibliometry will be an artefact in the future depends on the outcome of this battle. Here, we review the current status of this siege, how it arose and how it is likely to evolve.

KEY WORDS: Academic journals · Power law · Biomedicine · Bibliometric indices · Quality · Evaluation · Open access · Publishing model

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

INTRODUCTION

The advancement of science is based on a process of accumulating knowledge. Each new experiment or theoretical idea is always built in part on the results of previous research studies. Thus, essential to the evolution of scientific knowledge is an efficient system of communicating results and ideas within the academic community.

Scientists and the results that they produce differ greatly with respect to originality and impact within their own fields of study and on society at large. This calls for a system of evaluation when it comes to choosing which paper to read, to cite, or which person to hire in a research center.

The traditional printed journals have been the primary means of communicating research results, and as such have performed an invaluable service. They have

*Email: michael@damir.iem.csic.es

also served as an evaluation system since, by virtue of a thorough and objective peer review process by the scientists involved with them, they have, by and large, helped to maintain high standards. Each academic field has its own key journals, and publishing in one of them has always been considered as an index of academic quality and prestige.

Things are changing quickly. There is a rapid growth in the number of new researchers, publications, scientific journals and academic fields each year, and hence also of scholarly literature and demand for access to it by an increasing number of institutes and universities. Publishing houses have realised that science is big business. As a result, purely commercial interests have gained sway over many of the journals that we depend on for research information. Maximizing profits has become the controlling goal; a system that should serve the academic community and the public is now at the mercy of corporate acquisitions and profit-oriented planners, while the smaller publishers and societies left out from this race are being forced to meet the same challenges. Disseminating scholarly research seems to have become something of an afterthought. More worryingly still, more and more authors are reporting on the rise of censorship created by such a concentration of power. Bauer (2004) states:

'Minority views on technical issues are largely absent from the public arena... Since corporate scientific organizations also control the funding of research, by denying funds for unorthodox work they function as research cartels as well as knowledge monopolies... What national and international organizations publicly proclaim as scientific information is not safeguarded by the traditional process of peer review... the media need to know about and have access to the whole spectrum of scientific opinion on the given issue... A constant dilemma for reporters is that they need access to sources, and if they publish material that casts doubt on the official view, they risk losing access to official sources... In the bygone era, trustworthy science depended on scientists doing the right thing even when that did not immediately serve their personal purposes. In the new era of corporate science, the desires of individuals to serve the public good do not suffice to ensure that corporate actions will serve the public good.'

This explosion of academic literature has also made the problem of evaluation of science even more difficult. Quantitative indices (based on citation rates) that have been proposed as measures of academic performance, although in common use, have proven to be inappropriate for the evaluation of individual scientists (Seglen 1997). A consequence of this is that new alternative measures have emerged. Recent technological developments in information exchange have made it possible for the scientific community to open new doors, and it is beginning once again to take control over its own product.

We will begin this review by stressing the problems associated with market forces taking control over and

commodifying scientific publishing. We will then examine the various bibliometric indices used for academic evaluation and we will show how they are participating in the fortification of a large commercial publisher monopoly. We will focus on the mis-use of these indices and their manipulation by authors, reviewers and editors, locked in a continuous battle for power, money and prestige. Finally, we will examine the new initiatives appearing on the academic horizon to surmount these problems, such as the development of new open access bibliometric indices and publishing models that offset the costs currently paid to journals by authors and readers to fund open access alternatives.

THE COMMERCIALISATION OF INFORMATION

The crisis

A wave of mergers in the publishing business has created giant firms with the power to extract ever higher journal prices from university libraries (Lipscomb 2001). A few commercial publishers discovered that the easiest way to increase profits was to raise subscription prices and, specifically, that the fattest profits came from raising library subscription prices aggressively and relentlessly. Institutional subscribers, accounting for the lion's share of the revenue supporting publication of journals in most fields, paid the price (reluctantly and with increasing difficulty) because their users demanded access. With one foot in the door, these few commercial publishers built substantial portfolios of journals. The purchase of Harcourt General by Reed Elsevier is a prime example. The resulting union — with the addition of the science, technical, and medical (STM) division of Harcourt — controls more than 1500 journals, calculated to include 34% of the mainstream biomedical journals tracked by the Institute for Scientific Information (ISI) (Malakoff 2000). The high profits from these journals have funded further acquisitions and consolidations among publishers, and often these profits are diverted out of scientific activities into unrelated lines of business in order to enhance shareholder value. Feder et al. (2005) state:

'Can you imagine a company that simultaneously promotes arms sales and publishes health journals? Well, you don't have to imagine such a company — because it exists. Indeed, it is Reed Elsevier, the world's largest publisher of scientific and medical journals. And the finest of Reed Elsevier's journals is *The Lancet*, the leading global health journal, which has been receiving much attention from the Pentagon for its important articles showing that death rates in Iraq are far above those admitted by the United States government.'

Not surprisingly, such things are not taken lightly by academics who resent journals investing the fruits of

their labour in criminal or unethical activities and making access financially prohibitive to themselves and their peers. This resentment initially took the form of scathing articles, such as Feder et al. (2005), that were soon followed by a 2 yr boycott of Reed-Elsevier journals by hundreds of academics, a petition with 1915 academic signatories (including Nobel prize-winners I. McEwan, A. S. Byatt and N. Hornby) published in the *Times* newspaper (Stafford et al. 2007), the selling of nearly £2 million worth of shares by the Joseph Rowntree Charitable Trust, as well as weekly vigils outside Reed-Elsevier's London headquarters (CAAT 2007). The result, unsurprisingly, was a public announcement by Reed-Elsevier CEO Sir Crispin Davis that they will exit the defence sector (Reed-Elsevier 2007). As we shall see, this rebellion, that has ethics at its heart, is just the tip of the iceberg of an academic revolt now taking place more widely.

The overall results of journal conglomerate mergers are clear: high prices, declining circulation, inefficient production schedules and shady side-investments, such as Elsevier's involvement in the arms trade. Most of all, libraries, being held ransom to journal subscription 'bundles', and their users (often the same authors without whom journals would cease to exist), are suffering the most. One of these bundles (also known as 'the Big Deal') is ScienceDirect which claims to have 'over 25 % of the world's science, technology and medicine full text and bibliographic information'—corresponding to some '6.75 million articles' (Russo 2006). As a result, libraries are continually being forced to cough up more money or cut their journal subscriptions. Furthermore, it is now acknowledged that the publishing mergers represent a threat to free expression by reducing outlets for authors and the number of titles published (McNulty 1979, Lipscomb 2001). It has really become a crisis.

Just how over-priced is 'over-priced'

Between 1986 and 1998 real prices of academic journals approximately doubled, while real library budgets for acquisitions of books and journals rose by only about 50%. During the same time interval, the number of academic journals published increased by 60% (Bergstrom 2001). Libraries, unable to keep pace with these rising costs, are cancelling journals and diverting increasing amounts of the acquisitions budget to cover the cost of those remaining. Between 1986 and 1998 libraries on average cut the number of books purchased by 26% and the number of journal subscriptions by 6% (Bergstrom 2001). Furthermore, a dramatic disparity emerged between the prices charged by for-profit publishers, such as Elsevier, Wiley and

Kluwer, and those charged by non-profit societies and university presses. This gap widened in the 1980s and further widened in the 1990s, so that the for-profit journals charge about 5 times as much per page and 15 times as much per citation as non-profit journals (Bergstrom & Bergstrom 2004).

In the face of rising journal costs, some institutions have looked to pressure commercial publishers to reduce their charges. At Stanford University, where the cost of journal subscriptions had risen 50% in 5 yr (Miller 2004), the library was encouraged to 'systematically drop journals that are unconscionably or disproportionately expensive or inflationary.' (Miller 2004).

An article published in the *Wall Street Journal* on the pressure faced by journals from the shift to e-publishing and the revolt by academics to over-pricing, recently reported that the median expenditure on periodicals per library, during the period 1986 to 2003, rose from US\$ 1.5 million to US\$ 5.4 million—some 13% above annual inflation (Wysocki 2005).

In response to this, mathematicians, for example, have been urged not to submit papers to, nor edit for, nor referee for overpriced journals (Kirby 1997). Until recently, journal prices were not disclosed (their costs are still not widely accessible) and it took a 10 yr lawsuit by the American Mathematical Society (AMS) to 'win' this right. The AMS provides up-to-date information on the rise of mathematics journal prices (Rehmann 2004). In addition, journal prices by sector are also available (www.journalprices.com). This legal victory has helped to enforce an increase in transparency on publishing economics. Scholars can now make up their own minds about which journals are 'too expensive'. Editorial boards are beginning to wise up too. On August 10, 2006, all the editors of Reed-Elsevier's mathematics journal *Topology* resigned to protest against the outrageous prices of their publisher. A set of guidelines for pricing is available at: www.journalprices.com.

The over-pricing of journals can also be seen in the business news. Morais (2002) in *Forbes* magazine wrote:

'If you are not a scientist or a lawyer, you might never guess which company is one of the world's biggest in online revenue. Ebay will haul in only \$1 billion this year. Amazon has \$3.5 billion in revenue but is still, famously, losing money. Outperforming them both is Reed Elsevier, the London-based publishing company. Of its \$8 billion in likely sales this year, \$1.5 billion will come from online delivery of data, and its operating margin on the internet is a fabulous 22.'

Information has become a new commodity in the internet age. However, new technologies and growing access to free information worldwide make it vulnerable to attack (Waltz 1998). Today, pirated music downloaded from the web ravages corporate profits in

the music industry. Tomorrow could be the publishing industry's turn. The bottom line is that it is expensive to produce a high quality journal, even if it is a nonprofit-making journal or an open access electronic journal (see section 'The birth of autonomous scientific publishing'). What is clear is that a radical shift in the costing of journals is necessary if they are to survive.

The paradox

Academics are caught in a paradox. The advent of new technologies means that it has never been so easy and convenient to produce electronic copy as it is today (thanks to the internet and open source software such as TeX). It is possible to produce high quality online journals such as the *Annals of Mathematics* (published with the financial cooperation of Princeton University and the Institute for Advanced Study) for as little as 15 US Cents per page or less (Rehmann 2004). However, many key commercial journals are currently charging more than 4 US\$ per page of paper or electronic copy (Rehmann 2004; see also www.journalprices.com).

Scientists are writing the content and have developed, and are maintaining, the tools (like TeX) to format it and to provide camera-ready-copy (CRC). They often do peer review for free, give their work to the publishers for free and then buy it back from them at ridiculous prices. In an open letter to US university presidents and provosts concerning increasingly expensive journals Bergstrom & McAfee (1997) introduced the radical concept of overhead costs associated with peer review:

'So far, universities have failed to use one of the most powerful tools that they possess: charging for their valuable inputs. Journal editing uses a great deal of professorial and staff time, as well as supplies, office space and computers, all provided by universities. In any other business, these inputs would be priced.... However, we see no reason for universities to subsidize editorial inputs to journals that are priced to extract maximum revenue from the academic community.... We recommend the following policies:

- (i) Universities should assess overhead charges for the support services of editors working for journals that have basic library subscription rates of more than a threshold level of cost per measured unit of product.
- (ii) University libraries should refrain from buying bundled packages from large commercial publishers and should set clear minimal standards of cost-effectiveness for individual journals to which they subscribe.

We believe that it is reasonable to figure that a journal editor who handles about 100 papers annually would use about 20% of a secretary along with the associated space and other overhead materials, an overhead charge of at least \$12,000 per year would seem appropriate.'

As we shall see in the section 'Journals go on the offensive', the paradox means that such initiatives are

becoming more commonplace. Academics are sandwiched between the unwillingness or inability of their host institutions to pay publishing fees on the one hand and the ever-increasing competition to publish in the top journals on the other. Ironically, the paradox is also being intensified by authors having to pay for open access in both commercial and nonprofit-making journals (see section 'The birth of autonomous scientific publishing').

However, market forces alone are not solely responsible for the economic crisis in academia. We will demonstrate how the use and abuse of bibliometrics has played into the hands of the journal monopoly and policy-makers, effectively exerting a stranglehold on science and the acquisition and dissemination of knowledge. In the next section we will highlight the rise of the use of bibliometric indices, such as the journal impact factor, and we will assess their use in evaluation.

THE EVALUATION OF SCIENCE AND SCIENTISTS

The history of ISI and the Journal Impact Factor (JIF)

Scientometrics probably began with peer-to-peer scholarly letter-writing in the 17th century (Guédon 2002), but it came into its own in the 1950s and the beginning of the 'publish-or-perish' era. Moed (2005) states:

'It was no longer enough just to publish in bulk: it had to be demonstrable that your publications were also heavily used, hence useful and important. A direct indicator of usage was the fact that your research was cited by subsequent research.'

Eugene Garfield, the founder of ISI (now part of Thomson Scientific, a large worldwide USA-based publisher), originally put forward the idea of a citation index (a list of papers along with the papers that cite them) for the sciences as a way to improve the scholarly process (note that Gross & Gross [1927] had already proposed counting citations as a means of collection management for journals and law reports). Garfield (1955) also suggested the possibility of using citations as a measure of the impact of an article within its research field, meaning that the impact factor as a concept was first introduced over 50 yr ago:

'When one is trying to evaluate the significance of a particular work and its impact on the literature and thinking of the period . . . such an 'impact factor' may be much more indicative than an absolute count of the number of a scientist's publications.'

The total number of papers and the total number of citations comprise the pre-impact factor bibliometric indicators that were superseded on the grounds that

the former does not account for the quality of scientific publications, while the latter is disproportionately affected by participation in a single publication of major influence. As we shall see, a lack of correlation between quality and indicators keeps re-surfacing to haunt all bibliometric indices that have been proposed.

JIFs are calculated each year by Thomson Scientific for those journals which it indexes in its Journal Citation Reports (JCR) as follows:

$$\text{JIF}_t = \frac{C_{t-2}}{P_{t-2}} \quad (1)$$

where JIF_t is the journal impact factor for year t (e.g. 2000), C_{t-2} is the total number of citations to papers published in the journal in the previous 2 yr (e.g. 1998 and 1999) and P_{t-2} is the total number of papers published in that journal in those 2 preceding years.

Thomson Scientific also produces other statistics, such as (1) the immediacy index (the average citation number of an article during the first year of publication; (2) the journal cited half-life (the median time for articles cited in Journal Citation Reports to acquire half of all their citations); and (3) the aggregate impact factor for a subject category. The Impact Factor is, however, the most widespread of these metrics and according to Charles Jennings, editor of *Nature Neuroscience*, its release by the ISI every September 'triggers elation or gloom in editorial offices around the world' (Jennings 1998). The reason editors have such great expectations is because, while Garfield originally introduced the JIF as a means of evaluating research, it has mostly been used as a method of comparing the importance of journals (Brody 2006) — with journals, in turn, competing for library subscriptions largely based on their JIF (Garfield 1972).

Criticisms of the ISI and the JIF

Given the availability of citation indices for so many academic journals, it became easy for them to be used for the evaluation of individual scientists and research groups as well. There are numerous reports from many countries, such as Italy (Calza & Garbisa 1995), the Nordic countries (Hansen & Jorgensen 1995), Canada (Taubes 1993), Hungary (Vinkler 1986), Spain (Jiménez-Contreras et al. 2002), Japan and the US (Jennings 1998), where JIFs are used for the evaluation of individuals and institutions, or are being considered as a criterion for the allocation of university resources and positions. There are, however, many serious problems related to the use of the impact factor as a measure of scientific quality, either for individual researchers or for academic journals in general. Seglen (1997) has presented a brief and concise

summary of these problems. Below, we analyze the use and abuse of the JIF.

The JIF does not reflect the impact of the average journal article

The use of the JIF for the evaluation of individual academic performance is based on the assumption that it is representative of the average article published in that journal. This would be the case if the citation rate of individual articles in the journal showed a narrow normal (Gaussian) distribution around a mean value. However, evidence from various studies (Seglen 1992 and references therein) shows that plotting citation frequency against individual papers published in any journal (high or low JIF) results in a highly skewed and non-normal distribution with most citations coming from a tiny minority of the articles. The flipside to this is that the majority of papers are generally cited many times less (if at all) than the JIF suggests. This fact is brilliantly reflected in an editorial report in *Nature* (Nature 2005; see also Campbell 2008, this Theme Section):

'We have analysed the citations of individual papers in *Nature* and found that 89% of last year's figure [JIF] was generated by just 25% of our papers. The most cited *Nature* paper from 2002–03 was the mouse genome, published in December 2002. That paper represents the culmination of a great enterprise, but is inevitably an important point of reference rather than an expression of unusually deep mechanistic insight. So far it has received more than 1,000 citations. Within the measurement year of 2004 alone, it received 522 citations. Our next most cited paper from 2002–03 (concerning the functional organization of the yeast proteome) received 351 citations that year. Only 50 out of the roughly 1,800 citable items published in those two years received more than 100 citations in 2004. The great majority of our papers received fewer than 20 citations.'

In Fig. 1, we can clearly see that 88% of all citations come from just 50% of the articles. Worse still, half of the citations come from just 15% of the journal's articles. What all this means is that the JIF should apply only to journals and not to articles or authors, since it conceals information about individual papers. Eugene Garfield himself warns that the JIF would lead to a 'misuse in evaluating individuals' because there is 'a wide variation from article to article within a single journal' (Garfield 1998). Recently, Garfield (2005) likened his creation to that of nuclear energy:

'The impact factor is a mixed blessing. I expected it to be used constructively while recognizing that in the wrong hands it might be abused.'

In the section 'The actual distribution of JIFs', we offer an explanation for why there is such a non-normal distribution of citations.

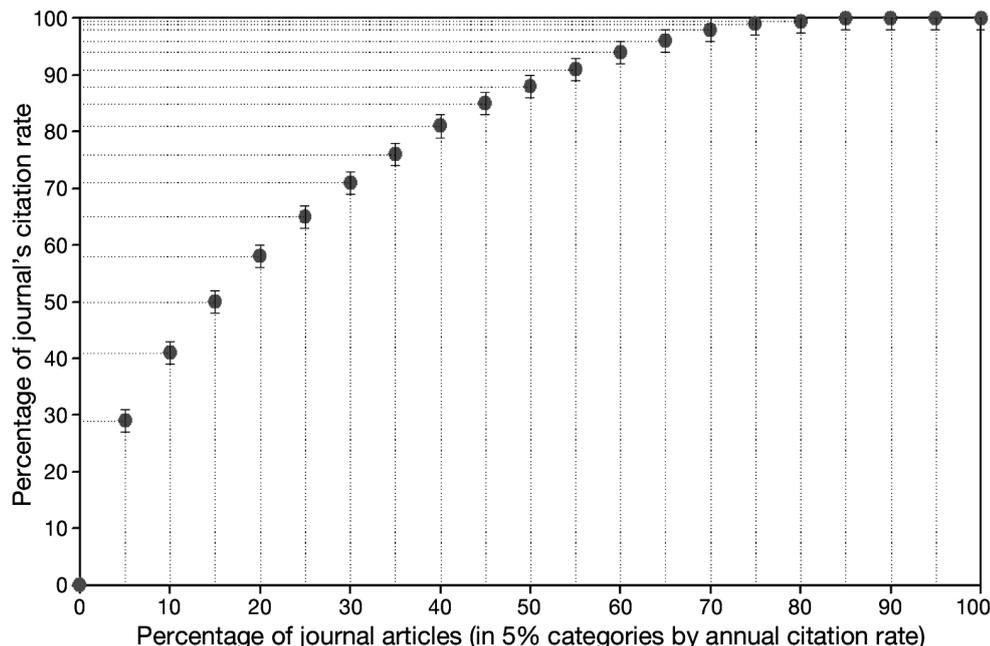


Fig. 1. Cumulative contribution of articles with different citation rates (from the least cited 5% to total journal impact, 100%), reconstructed from average values for 3 biomedical journals (data taken from Seglen 1992)

Other biases that affect the JIF

The JIF is equal to the number of citations from articles in any given year to items published in the journal during the 2 preceding years, divided by the number of such items. ISI, however, is not dividing like with like. It is including in the denominator refereed articles, notes and reviews, while in the numerator more types of document are included (also encompassing editorials, letters to the editor and conference proceedings, to name a few extras that also contain citations) (Seglen 1992). This bias inflates the fraction and hence increases the JIF. In the section 'Biomedicine: at the rough end of the journal monopoly' we will discuss how these vulnerabilities in the calculation of the impact factor are being exploited by editorial boards and authors for commercial reasons, particularly in the field of biomedicine, in order to increase the JIF for their own ends.

Another bias inherent to the way the JIF is calculated has to do with the 2 yr time window. For example, in social sciences, it usually takes longer for an article to receive a meaningful number of citations, while in rapidly evolving scientific areas like biomedicine, citations accrue much faster. Thus, although the time window used in the calculation of the impact factor may be appropriate for some fields of science, such as molecular biology, it is not appropriate for subjects with a slower publication pattern, such as ecology (Seglen 1992).

The absolute number of researchers, the average number of authors on each paper, variations in citation habits between different disciplines and the nature of results in different research areas all combine to make the

JIFs a very inappropriate bibliometric for the comparison of journals from different subject areas or different types of journals (e.g. review journals vs. experimental journals) (Fassoulaki et al. 2000). It is well known that medical journals have higher JIFs than mathematical and engineering journals. In the context of competition for state or even regional funding, scientists from the latter are often left in the academic wilderness.

ISI provides citation data for approximately 8700 journals of an estimated 24 000 in circulation (Harnad et al. 2004). ISI justifies this under-representation by a principle referred to as Bradford's law; this claims that 'the core literature for any given scientific discipline is composed of fewer than 1000 journals' (Garfield 1979). However, this jeopardizes global academic coverage.

In addition to including only a small percentage of all available journals in its database, ISI has also been criticised for having a low 'international coverage' (Brody 2006). Although ISI's Web of Knowledge indexes journals from 60 countries, the coverage is very uneven. Very few publications in languages other than English are included and very few journals are included that are produced by less-developed countries (Buela-Casal et al. 2006). The bias towards English language journals means that the non-English journals included in JCR have lower JIFs since most citations to papers in languages other than English are from other papers in that same language. Thus, the relatively higher (apparent) academic impact of American science represented by the JIF is, to an important extent, determined by its greater share value of publications, self citations, and citation biases (Seglen

1997). Moreover, the exclusion of local journals from ISI's databases produces an underrepresentation of scientific output, which can have important ethical and science policy implications, especially in disciplines that benefit greatly from area-specific information, such as ecology (Stergiou & Tsikliras 2006).

Are the alternative single parameter measures any better?

A host of other citation indices have been proposed to supplement or complement the JIF. These include (1) the PageRank (Bollen et al. 2005), the Eigenfactor (Bergstrom 2007), the h-index (Hirsch 2005), the g-index, (Egghe 2006), the v-index (Vaidya 2005), the Euro-Factor (EF)TM (Hofbauer et al. 2002), and the Reading Factor (Darmoni et al. 2002). Since we have devoted some attention to the limits of the JIF, we will attempt the same for these other measures.

A plot of the PageRank (or the Eigenfactor, which also uses the page rank algorithm) versus JIF (Bollen et al. 2006, their Fig. 1) reveals a strong non-linear relationship between the two and a disappointingly large scatter over the whole range of impact factors. The h-index, although reflecting the modal average number of citations for a scientist (i.e. it is statistically insensitive to outliers—papers having a disproportionate number of citations), prejudices against scientists with short academic careers or who have articles published in low JIF journals where they are not often cited (if at all). For example, the brilliantly influential Évariste Galois' (who died young having published only 2 papers) has an extremely low h-index of 2 that will remain the same forever, independently of the number of citations that these 2 articles receive in the future.

As with all citation-based enquiries and indices (this applies to the remaining citation indices listed above that are variations on the JIF), authors who have published prior to automated indexing that began in 1989 are likely to have underestimated values. Finally, while quantitative indices are intended to measure scientific productivity, there is a danger in reducing human activity as complex as the formal acquisition of knowledge down to a single numeric metric (Kelly & Jennions 2006). Without exception, the result of all of these single-parameter measures is the same: pressure on scientists to publish in the best journals, and pressure on universities to ensure access to that core set of journals.

The problem of quality

A general criticism of the JIF is that it is being wrongly used as a metric of quality. Quality is a more

complicated concept than impact assessed through citations, and it cannot possibly be measured by a single-parameter index like the JIF (Kelly & Jennions 2006). The American Society for Quality (www.asq.org/) defines quality as, 'a subjective term for which each person has his or her own interpretation' and, in Pirsig's *Zen and the art of motorcycle maintenance*, quality is defined as the 'union of aestheticism with functionality' (Pirsig 1974). Most people would agree that there is a strong subjective element to quality.

In previous work (Buela-Casal et al. 2006), in an attempt to improve our understanding on the quantification of subjective concepts, we performed a quantitative analysis of the (simpler) concept 'internationality'. We found that single-criteria measures of internationality gave inconsistent results and were incapable of distinguishing between the internationality of a sample of journals. We drew the conclusion that internationality is more adequately described as being 'a matter of degree' and is better assessed using a multi-parameter approach (Perakakis et al. 2006). In a similar manner, we believe that the assessment of quality should be approached by multi-parameter indices that can better describe our daily judgements and subjective perceptions.

Associating citation counts alone with quality means ignoring all other parameters that may be relevant to its measurement. In a similar vein, the idea of multi-parameter indices to assess academic performance has been applied by Index CopernicusTM International, a subscription-based research infrastructure providing information and tools for scientists, research administrators and government agencies. Quoting text from www.indexcopernicus.com:

'Based on our US patent pending mathematical formulas, the Index CopernicusTM (IC) also provides an also interactive and reliable scientists' evaluation system. IC performs multi-parameter analysis of scientific output and research potential, both of individual scientists and research institutions.'

Unfortunately, however, the black-box nature of the quality measurement method, guaranteed by its pending US patent, means that its scientific basis cannot, at least for the moment, be assessed. The absence of published journal articles on the quality assessment algorithm further frustrates this effort. For the present at least, we are neither in a position to objectively review it nor to pursue it further. We encourage others to participate in the development of multi-parameter indices which are, to date, lacking from the literature.

A new development from Stevan Harnad has just come to light at the time of writing. At the 11th Annual Meeting of the International Society for Scientometrics and Informetrics in Madrid, 25 to 27 June 2007, and in Harnad (2007) he explained that:

'The first thing psychometricians would do with a 'battery' of univariate metrics would be to systematically validate them against external criteria that already have some face validity for us. There are other classic performance measures, such as funding, doctoral student counts, and prizes—but, frankly, using those would be circular, as they have not been externally validated either. What psychometricians sometimes do first with their batteries of diverse metrics is to look at their inter-correlational structure through principal-component and factor analyses.'

We may be criticised for 'post-bibliometric thinking' by bringing subjectivity back into the picture, but the next section reinforces this point.

Nobel Prize winners and the ISI Nobel Laureate Predictor failure

For many, a Nobel Prize is the ultimate measure of quality. Indeed, Thomson Scientific's Scientific Laureates was launched with this in mind (<http://scientific.thomson.com/press/2006/8340634/>):

'Citations are an acknowledgement of intellectual debt—a direct demonstration of influence in a given subject area... Over the past 30 years, our studies have demonstrated a strong relationship between journal article citations and peer esteem... Thomson Scientific is the only organization to use quantitative data to make annual predictions of Nobel Prize winners.'

However, the 4 cases of 'correct' predictions by Thomson Scientific based on citation counts show no correlation at all with citation rank. Andrew Z. Fire, who won half of the prize in biology and medicine for the discovery of RNA interference, is citation-ranked 222nd in molecular biology and genetics. He shared the prize with Craig C. Mello ranked 392nd. The third Scientific Laureate, Roger D. Kornberg, who won the prize in chemistry for his studies of the molecular basis of eukaryotic transcription, is not ranked but was quoted as having, '6 highly cited papers of the last 10 years'. Finally, the physics prize—shared by John C. Mather (with no reference to citations or rank) and George F. Smoot, ranked 220th in physics, for their discovery of the blackbody form and anisotropy of the cosmic microwave background radiation—confirms the anti-correlation. In addition, there have been many published cases of scientists being rejected by high impact factor journals only later to win Nobel Prizes for those same articles (Campanario 2003). Clearly then, citation ranks, Nobel Prizes and the perception of quality do not correlate so simply.

The problem of evaluating science and scientists does not end here. In the bibliometric analysis of highly skewed distributions, such as the one we have already met for the citations within an individual

journal, bad statistics also raises its ugly head, but for another more complicated reason. Here to a certain extent, bibliometricians can be forgiven. The reason in part is due to the complexity of the mathematics of such distributions (known as 'power laws') and the very recent understanding of the mechanisms that give rise to them. It is precisely here that we will find the explanation for how the use and abuse of bibliometrics has played into the hands of the journal monopoly.

BIBLIOMETRICS AND THE JOURNAL MONOPOLY

The picture arising from the non-normal power law distribution of the citation rates of individual articles in a journal can be visualized, in the words of Lehmann et al. (2005), as:

'a small number of active, highly cited papers embedded in a sea of inactive and uncited papers.'

However, citation rates are not the only thing that follow this type of distribution; JIFs do also. Garfield and Sher were the first to present research into the citation behaviour of academic journals (Garfield & Sher 1963). They found that when plotting citation frequency for research literature published during 1961 only a small subset of journals received the majority of citations: just 60 of the 5000 journals analysed accounted for 60 % of all citations. Years later, Garfield (1990) pointed out that:

'No matter how many journals are in the market, only a small fraction account for most of the articles that are published and cited in a given year.'

Thomson ISI (2004) estimates that 'a core of approximately 2000 journals now accounts for about 85 % of published articles and 95 % of cited articles.' Although this inequality in the academic journal landscape that we call 'the journal monopoly' has been known since the 1960s, recent developments in network theory now allow us to make mathematically testable hypotheses about the mechanisms generating such monopolies. We will show (see section 'Biomedicine: at the rough end of the journal monopoly') that the bibliometric abuse enshrined in the JIF not only accentuates the journal monopoly, but may also be responsible for its coming into existence in the first place. In order to understand this, we need a little theory from bibliometrics.

Bibliometrics is the quantitative analysis of publication patterns within a given field or body of literature. Three commonly cited laws of bibliometrics all have non-normal distributions like the distribution of citations in a journal: Lotka's Law (Lotka 1926), Bradford's Law (Bradford 1934) and Zipf's Law (Zipf 1949). Of particular relevance to the distribution of JIFs and citations is Zipf's Law:

$$P_n = cn^{-a} \tag{2}$$

where P_n is the frequency of occurrence of the n th ranked journal, c is a constant and a is the exponent which has a value close to 1. When plotted on a log-log graph this equation follows a straight line of slope equal to the exponent (or 'power') a . The distribution of papers by citation impact has been said to adhere to the Zipfian power law (Redner 1998). However, in the next section we will show that this is not entirely true.

The actual distribution of JIFs

Lavalette (1996) extended Zipf's law with an excellent fitting to data for JIFs. It is barely more complex than Zipf's law, and has the form,

$$P_n \approx \left[n \frac{N}{(N-n+1)} \right]^{-a} \tag{3}$$

with the role of n as independent variable being replaced by the ratio $n/(N-n+1)$. In the context of scientific journals, P_n is the JIF that equals the frequency of a set of N journals, ordered by the descending ranking number n with 2 fitting parameters—the exponent a and a scaling constant of proportionality.

For a sample of 8011 journals between 1974 and 2001 (the JIFs of which were kindly provided by Ioan Popescu), we calculated the Lavalette fitting and its deviation from the Zipf Law which is shown in Fig. 2. From the 23rd highest JIF-ranked journal to the 1691st journal, the data follows almost perfectly a straight-line Zipf Law. The existence of this power law means that the top 1691 journals follow a strict hierarchy, with the JIF falling dramatically as we move to lower ranked journals. The Lavalette Law also fits this region of the graph, but also fits the tail in the data from the 1691st journal to the lowest JIF ranked 8011th journal. To understand these trends, as well as the 'King Effect' for the very top 23 JIF ranked journals, we hypothesise that the answer comes from the theory of social networks. In essence, 'preferential attachment' (Barabasi & Albert 1999) is governing the distribution of JIFs, i.e. it may be the underlying mechanism responsible for the journal hierarchy (for a review of the empirical evidence for such mechanisms, see Newman 2005). Picturing the journal system as a network of interconnected nodes with journals as nodes and author citations linking them together, preferential attachment means that new citations tend to cite those journals that have already attracted many citations. In other words, some of the

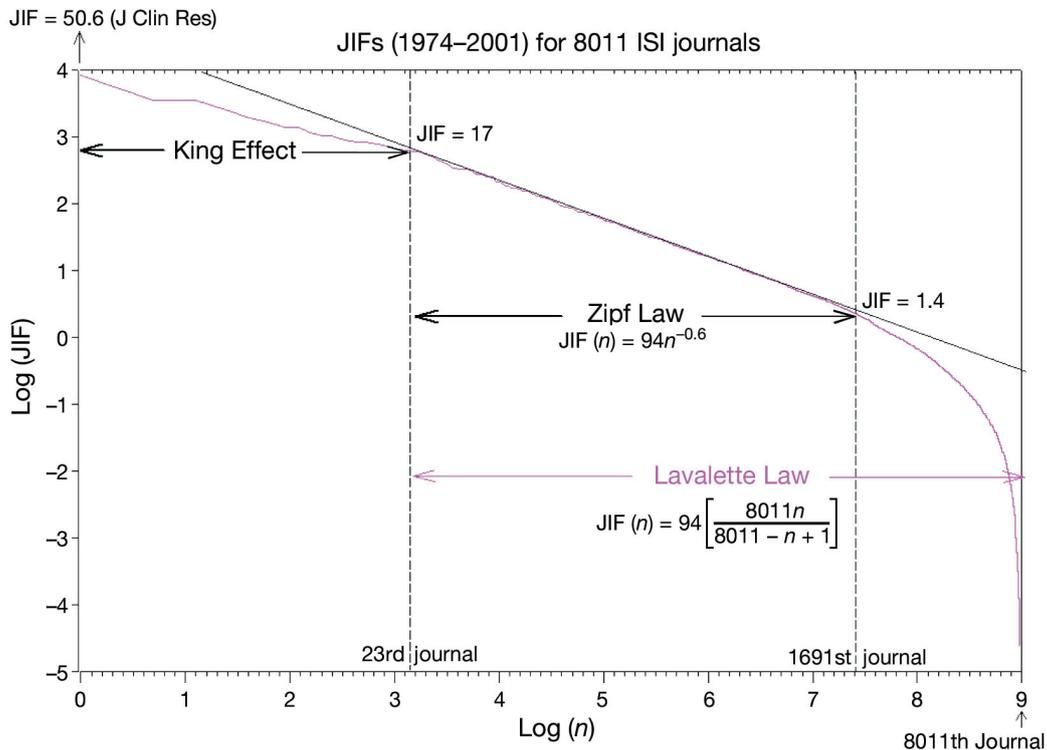


Fig. 2. Rank-frequency log plot of journal impact factors (JIFs) (vertical axis) for the 8011 Institute for Scientific Information (ISI) journals (horizontal axis) having calculated impact factors between the years 1974 and 2001. The black line is the Zipf Law fit to the data between the 23rd journal (JIF = 17.0) and the 1691st journal (JIF = 1.4). The pink line is the Lavalette Law fit between the 23rd journal (JIF = 17.0) and the 8011th journal (JIF = 0.125). The 'King Effect' for the top 23 journals clearly deviates from the Zipf Law

journals act as highly cited 'hubs' with the majority of them having few or no links (cites). A highly cited journal (or article) is more likely to receive more citations, since it is already more visible and prestigious. Preferential attachment produces Zipf Law distributions (like the one in Fig. 2) in many other networks, such as the internet (Faloutsos et al. 1999), the web of human sexual contacts (Liljeros et al. 2001) and the protein network in the human cell (Jeong et al. 2001). This means that authors for the 1668 journals in this part of Fig. 2 are preferentially attaching to journals in direct proportion to their JIFs, i.e. they prefer to cite journals with high JIFs. To our knowledge, no other mechanism explaining the rise of the 'journal monopoly' has been provided. This leaves 2 other trends to be explained.

The first is the deviation from Zipf's Law for the top 23 journals: the King Effect (Laherrere & Sornette 1998). In network theory, the 'king' or top part of the hierarchy has so much 'power' that it can override the preferential attachment law. In the case of academic journals this modulation effect suppressing Zipf's Law, we believe, comes from disproportionately high rates of rejection and maybe lower visibility (therefore lower citations) due to excessive access fees. Prestigious journals, besides being expensive, base their reputation on their exclusivity, and a 'top-down' approach, creaming off a tiny proportion of the articles they receive. Rejection rates at *The Lancet* and the *New England Journal of Medicine*, for example, are well over 90% (Wager et al. 2006).

The second is the striking downward deviation from a Zipfian straight line for the lowest JIF ranked journals running from the 1692nd to the 8011th journals in the sample (Fig. 2). Zipf's Law disappears entirely, indicating a dramatic change in preferential attachment for the lowest impact factor journals. We believe that due to their lower values of JIF (below 1.4), authors' citing behaviour is now not ruled by preferential attachment. Hence, the selection effect is lost and instead authors are now citing articles based on criteria other than JIFs, such as relevance of content.

Network theory has helped us to realise that the mechanism responsible for the growth of the journal monopoly appears to be preferential attachment of authors to certain journal-hubs under the inertial pressure of inflating their 'personal impact factor'. The abuses of the JIF described earlier, when combined with this 'rich get richer' positive feedback loop, mean that bibliometrics alone seems to be fortifying the journal monopoly. Lack of awareness of this, or (worse still) academics playing along due to their perception that there is no alternative, yields acceptance of the 'publish or perish' dogma, leading to great competitive stress.

With biomedicine occupying the top positions in this hierarchy, power politics have come into play. This leads to a rigged publishing game between authors, editors and reviewers, as we explain in the next section.

BIOMEDICINE: AT THE ROUGH END OF THE JOURNAL MONOPOLY

Procrustean logic

It has been acknowledged that scientific innovation has prospered most when talent, supportive institutions, mobility, free communication, and finance are available in significant measure (Hall 1998). Since the topic of this review is the siege of science, then the funding of scientists and scientific institutions is of particular relevance. Since the Second World War, biomedical research has been the beneficiary of parallel advances in the physical, social and information sciences. That momentum greatly expanded financial support for biomedical research, which in the early 1980s started to exceed the funding of engineering and the physical sciences (US Census Bureau 1999).

In the USA, biomedical research support comes from the following 4 major sponsors: (1) federal government, (2) state and local governments, (3) private not-for-profit entities including foundations and (4) industry. Biomedical research funding increased from US\$ 37.1 billion in 1994 to US\$ 94.3 billion in 2003. Notably, the principal research sponsor in 2003 was industry (57%) followed by the National Institute of Health (28%) (Moses et al. 2005). In Canada and Europe, as in the USA, the line between research and commercial activities became increasingly blurred, heightening public sensitivity to potential conflicts of interest in biomedical research. In November 2005, a panel of medical experts commissioned by the Canadian Association of University Teachers warned that outside pressures were putting the integrity and independence of as many as 20 000 clinical faculty staff and researchers in Canada at risk (Payne 2004). In Europe the commercialisation of research is also causing increased concern. 'Universities are being encouraged to make commercial use of their discoveries,' says Imogen Evans, research strategy manager for clinical trials at the Medical Research Council in London; 'This is blurring the once clearly strict dividing lines between universities and commerce' (Agres 2005)

Making a product out of scientific results not only raises questions over integrity and freedom of research: it also puts pressure on areas of knowledge or other scientific activities that do not have the clear short-term commercial relevance that biomedicine has

for the corporations. The humanities, for example, are suffering more and more from frozen or reduced budgets, due to the money flowing towards more sponsor-attractive areas (Sutherland 1999).

This commercial relevance of biomedicine is also translatable into bibliometrical terms. The JCR for 2006 lists 6120 journals spanning all fields with a JIF above zero. Of these journals, only around half (3370) have a JIF above 1, and very few have a 'high' JIF. It is really effortless to see the correlation between the journals with the highest JIF and the most funded area of scientific research (biomedicine), as shown in Table 1.

Only 1.7% of all journals have a high JIF (which we define to be above 10), and of these only 21 (of the 109) are not biomedical (Table 1). Does this mean that biomedical research is to be evaluated or considered more important than other areas of research? Not at all. The scientific field to which a journal belongs greatly influences its JIF (Scully & Lodge 2005). For instance, the highest impact factor in the ISI subject category 'Chemistry' is 5.6, whereas that in 'Cell Biology' is a huge 31.3. This cannot mean that Cell Biology is 6 times more important than Chemistry. Different disciplines have widely differing citation practices, and to rank a discipline based on the JIF of its journals is to make a determination based on Procrustean standards (Hecht et al. 1998). Because of this, Linardi et al. (1996) suggested that comparisons of the journals on the basis of the JIF should be limited solely to intra-area evaluation since 'inter-area comparisons may be both inappropriate and misleading' (Linardi et al. 1996). Furthermore, the practice in relation to the comparison of journals using the JIF between different subject areas is distinct: in areas like science or medicine, publications in high JIF journals are the currency by which professional worth is measured. In a tenure review, or job interview 'referenced articles' count heavily. In humanities subjects (English, history, modern languages) they do not carry the same weight (Sutherland 1999). For example, the Australian Council for Humanities, Arts and Social Sciences stated that, 'standard bibliometric practices do not capture the variety of research outputs in the humanities and social sciences and are clearly absurd in the creative arts' (Sutherland 1999). For this reason, it has been proposed that, in order to evaluate research

journals in these areas, several indicators should be used: level of holdings in large overseas academic libraries, coverage in international databases, standards of refereeing, peer evaluation by researchers and frequency of citations in the ISI (East 2006).

Moreover, the logic 'the journal is more important than the scientific message' is having an even more devastating effect on science in developing countries, where money for science is much scarcer than in developed countries. However, the 'tyranny' of numerical assessment of scientific merit reduces the possibility of defining independent locally oriented scientific policy (Barcinski 2003). Funding agencies and the ISI are defining the dimensions for the modern but similarly cruel 'scientific procrustean bed'.

Playing the game

Many people believe that the battle for publication in leading journals in the area of biomedicine is becoming brutal. The 'publish or perish' directive imposed by the funding agencies is becoming almost the sole rule in this vicious game. Senior scientists have to publish their results in 'top journals' in order to get grants. Research committees are investing in young scientists only when they have papers in leading journals as a guarantee of future fundability (Insall 2003). Given this, there is an increasing concern about the use of the JIF by authors, referees and editors.

The editors

The editors' main concern is to maintain or to increase the high JIF of their journal. This objective has consequences in terms of what an editor's job involves nowadays. *Nature* receives around 9000 manuscripts a year and rejects about 95% of biomedical papers (2003 figures). The remaining 5% should be the 'trendiest' papers of highly evaluated (which usually means highly cited) authors that the editors believe to have a higher probability of increasing their journal's JIF. Papers by Nobel Prize winners, for example, can receive 30 to 40 times more citations than the average article (Sher & Garfield 1966).

Moreover, certain technicalities can be used in order to manipulate the JIF. Editors can choose to publish more review articles. While many research articles remain uncited after 3 yr, nearly all review articles receive at least 1 citation within 3 yr of publication—therefore raising the numerator. Furthermore, the length of review articles means that fewer overall articles will be published in a journal, reducing the denominator. The combined effect is there-

Table 1. Link between biomedical journals and high journal impact factor (JIF). Note that of the 25 journals with the highest JIF, only 2 are not biomedical

High JIF	Number of journals	Not biomedical
63.33 > JIF > 22.9	25	2
JIF > 10	109	21

fore a substantial inflation of the JIF (Seglen 1997). Although nobody would deny that review articles are important distillers of knowledge, one can see how this could be used as an inflationary tactic. An editor can publish a large correspondence section, as well as controversial editorials, as the citations that these items receive contribute to the numerator, whereas letters, editorials and abstracts from meetings are not counted in the denominator and therefore help inflate the JIF (Scully & Lodge 2005). Interestingly, the acceptance of papers of controversial scientific content may also increase the JIF, as they are likely to be highly cited (even though they may be the subjects of criticism). Although Garfield (1979) stated that negative citations are of limited importance, the fact that there is no discrimination in the calculation of the JIF between positive or negative citations creates an inaccuracy in the calculation of article quality based on the number of citations alone (Moed & Van Leeuwen 1995).

It is obvious that this evaluation pressure has led editors to find ways to cheat the system in order to maximise the JIF. Another mechanism, which may have important ethical considerations, is that some editors encourage authors to add references to articles published in their own journal. Fassoulaki et al. (2000) found a significant correlation ($r = 0.899$) between self-citing rates and JIF. However, it should be noted that it is almost universal for articles in a journal to cite primarily its own articles since they are considered to be of equal merit in the same specialised field.

The authors

Funding agencies and promotion committees feel obliged to evaluate scientists, but they often lack the detailed knowledge that is necessary to assess the quality of scientific work. Furthermore, peer review by competing experts in the same field means that it is prone to bias due to human foibles. So, evaluation of scientists is currently based on the number of their publications, their position in the authors' list and the JIF of the journals in which they publish. Money and positions are distributed based on performance indicators (personal JIFs) that, in Spain and Japan for example, have reached formulaic precision (Lawrence 2003). Thus, authors are expected to play the game that is defined by the JIF the best way they can, and of course this does not necessarily mean doing the best science they can.

Playing this game, where authors are under intense pressure, has led to a series of tactics (unconnected to the scientific work done) that has affected the process of writing and submission of results. Since the editors

are the first people that authors need to convince, courting them is expected, as is networking at conferences, using contacts or even issuing threats (Lawrence 2003). Editors will also need convincing to send the paper out for peer review, as only a minority of papers in the 'top' journals make it. Authors make 'bold claims' about the 'extraordinary novelty of results', not only to 'trick' the editor into sending the article out for peer review, but also knowing that the reviewers will then later ask for such hyperboles to be removed. Many worry that this is a dishonest approach that is rewarded by the current system (Brookfield 2003).

We have already mentioned that the area of biomedicine is the one most strongly affected by the regime imposed by the JIF. Scientists in this field often have interesting basic research results that are dressed up with links (however tenuous) to human disease in order to make them look more attractive (Lawrence 2003). Even if that link is vague or irrelevant it can raise the number of citations that help the author to get an important grant from the pharmaceutical corporations.

When submitting their paper, authors often pay careful attention to timing, since the fashions in science can also affect the fate of their article. Editors sometimes choose not to publish an interesting article only because it is not 'fashionable'. Timing is also very important for the highly competitive laboratories; it is best to know when your competitor is planning to submit, so you will not be scooped. We stress that these problems are most acute in biomedicine, since this area dominates the top journals. These extremes are less likely to occur in journals of lower JIF (in the middle or the tail of our Zipf plot; Fig. 2).

Another issue is that both referees and editors have no way of screening for false authorship. This means that an author with minor (or no) contribution to the work can be added to the author's list (Brookfield 2003). The competition for positions or funding is so tough that young scientists, who desperately need publications in a 'top' journal for their careers, sometimes seek to win inclusion in the author list as a favour from their group leaders. We see, once again, how pressure to inflate personal performance indicators (that are based on JIFs) is strong in highly competitive areas such as biomedicine.

All cases of misuse of the JIF or those instances of abuse of the existing system may not be the norm, but all the facts point towards the same conclusion: the evaluation of science and scientists based on the JIF and other single-parameter indices is problematic, and further investigation needs to be done in order to design a better index to measure scientific quality.

The reviewers

The Editor of *The Lancet* (Horton 2000) wrote:

'But we know that the system of peer review is biased, unjust, unaccountable, incomplete, easily fixed, often insulting, usually ignorant, occasionally foolish, and frequently wrong.'

The future lives and careers of young authors are mainly dependent on the decisions of the reviewers that judge their papers. A fair reviewing system therefore is in evident need, but is not an easy task. Johannes Schmid in his article 'Cover the eyes of Lady Justice' (Schmid 2003) lists a series of issues related to reviewers and their complicated job: (1) reviewers are scientists and also authors, making them even more prone to conflicts of interest; (2) since they know the names and affiliations of the authors, it could be a tricky task to judge the article of a good friend or (even more difficult) of an author from a rival laboratory; (3) mutual favours do occur in reviewing that could result in the emergence of scientific lobbies and the establishment of 'aristocrats' in the scientific community (Schmid 2003); (4) scientific aristocracy is very difficult to enter into as a young non-established author; and (5) it is very hard to compete for space in a leading journal, since this scientific 'elite' has the power to define what is fashionable or 'hot' in a given subject area. Schmid (2003) proposes 'blind submission' as a guard against these dangers. We share his worries and have included 'blind submission' in the publishing model we propose (see section 'A blueprint for a "green" OA publishing model').

There is another very important issue that reflects how problematic the current peer review process can be. Several scientists complain of their manuscripts being stalled at the peer review stage, only to find that very similar findings suddenly appear in another journal. A referee might have discussed the paper at the review stage with colleagues who then abused that privileged information. The case of Paul Chu is well known. He was contacted by 2 other groups to tell him that the material he was describing in his paper to be used as a superconductor, was erroneous. This happened while his article was still under review (Dalton 2001). Even though the scientific community chooses to believe that these cases are rare, they keep appearing in the media — making it difficult to ignore them. The fact that this type of abuse has been related to journals of high esteem and JIF raises worries about the situation in journals that receive less public attention. This is a typical case of an abuse allowed by the monopoly established by the JIF.

All the concerns we have raised have created a critical mass pushing scientists over the edge. Discontent and survival are strong emotional forces that have dri-

ven scientists to invent and take advantage of the new technologies in order to create alternatives to the journal monopoly. In the next section, we will discuss scientists' attempts to make the publishing game fairer.

THE BIRTH OF AUTONOMOUS SCIENTIFIC PUBLISHING

Authors, unsatisfied with the measurement of quality as reflected by single-parameter indices, priced-out by journals, denied ownership and access to their work and that of their peers, are jumping ship in droves from expensive commercial publishers. With need as the mother of invention, the internet has allowed the development of new modes of electronic access, code-sharing protocols and publishing — with open accessibility as an underlying and guiding principle:

'If you have an apple and I have an apple and we exchange these apples then you and I will still each have one apple. But if you have an idea and I have an idea and we exchange these ideas, then each of us will have 2 ideas.' George Bernard Shaw (Delamothe & Smith 2001)

Open access (OA)

The first major international statement on open access from the Budapest Open Access Initiative (BOAI) (see Chan et al. 2002) declared:

'Open access (OA) is immediate, free and unrestricted online access to digital scholarly material, primarily peer-reviewed research articles in journals.'

This provided an initial definition, and the statement to date has a list of 4896 signatories. The Directory of Open Access Journals (DOAJ), where OA policy is based on this statement, now contains 2858 fully OA journals with 151 395 searchable OA articles (www.doaj.org; 6 October 2007).

OA research material can be roughly separated into 4 sources: OA journals, subject-based research repositories, institutional repositories and personal web sites. A scholarly search tool would ideally allow the user to search or navigate across all OA material, providing the most useful matches first. The Open Archives Initiative Protocol for Metadata (OAI-PMH) for OA journals and repositories has been developed for this purpose. It has achieved wide acceptance due to its exploitation of widely deployed web technologies such as HTTP and XML (Van de Sompel & Lagoze 2002). However, search and indexing services like OAI-PMH, although successful in retrieving OA articles on a large scale, still have barriers to surmount. The reason is that, more often than not, they force users to use a citation search tool such as SCOPUS that may lead cor-

rectly to the target paper, but then redirects authors via hyperlink to the journal that published the article. Hajjem et al. (2005) found that only 5 to 16% of papers indexed by the ISI Science Citation Index were available as OA, and so authors end up having to pay a subscription to view the full text of their citation search.

The justification for OA is that it allows users of academic literature to access the material they need to perform their research. Most research is published in journals or conference proceedings that are only accessible by paying a subscription fee and, with 2.5 million research papers being published annually (Harnad et al. 2004), no research library can afford to pay for access to all the world's research output. The situation for researchers in developing countries is worse (Chan & Kirsop 2001, Smith 2004a). As a result of such economic pressures, OA began to take root.

There are 2 main currents in the OA movement (Harnad et al. 2004): OA self-archiving ('the green road') and OA publishing ('the golden road').

OA self-archiving

OA self-archiving was first formally proposed by Stevan Harnad (Harnad 1995). Authors make unrefereed, pre-print versions of their articles freely accessible online by depositing them (without charge) in institutional or subject-based repositories, such as arXiv (<http://arxiv.org/>) and Citeseer (<http://citeseer.ist.psu.edu/>), or by housing them on personal websites. Archiving costs are currently covered by their host institution or the institutional owners of repositories (we will look in detail at costing models in sections 'The economics of open access' and 'In defence of public science'). Journals enter the picture if authors then decide to submit their articles to peer-reviewed journals. Harnad (2003) has proposed a method which he calls the 'preprint-corrige strategy' to get around journal copyright agreements. Authors, having received reviewers' comments, can then deposit a peer-reviewed copy with corrections in an OA repository. This means that articles can be deposited without authors having to pay publishing fees or being bound by copyright. In addition, scholarly standards can still be maintained due to peer review without costs to authors and readers.

However, the value added by the role traditionally played by professional copyeditors and typesetters working for journals is absent in this model. Author self-archiving, when combined with the harvesting power of OAI-PMH enables a parallel and 'free' (but see also section 'The economics of open access' on the costs associated with the green road) system of publishing to exist alongside the paid-for access one, without significantly changing the economics (Kyrillidou & Young 2004).

OA publishing

Authors can publish in OA journals (some of which take a fee from the author) that make their articles freely accessible to readers online after publication. For example, the DOAJ (www.doaj.org)—hosted, maintained and partly funded by Lund University Libraries Head Office, the Open Society Institute, SPARC, BIBSAM and Axiell—state:

'from the BOAI definition of "open access" we take the right of users to "read, download, copy, distribute, print, search, or link to the full texts of these articles" as mandatory for a journal to be included in the directory'

All content at DOAJ is freely available and without delay (no embargo period). However, the DOAJ discriminates between what it calls OA journals (no publication fee to authors) and hybrid journals which charge a publication fee to authors. The Public Library of Science (PLoS) is an important test case of such an OA publisher. The situation is worst in biomedicine where commercial publishers dominate the JIF hierarchy. PLoS was formed originally to fight this, and circulated an open letter calling on publishers to adopt reasonable policies. 30 000 scientists signed this letter and pledged to:

'publish in, edit or review for, and personally subscribe to only those scholarly and scientific journals that have agreed to grant unrestricted free distribution rights to any and all original research reports that they have published, through PubMed Central and similar online public resources, within 6 months of their initial publication date'
(PLoS, www.plos.org/)

Unsurprisingly, the response from commercial publishers was chilly. However, PLoS persevered and set up its own journals in biology and medicine with the help of a US\$ 9 million grant from the Gordon and Betty Moore Foundation. In 2007, PLoS Biology entered the bibliometric scene with a JIF of 13. However, when this grant was used up, PLoS had to start charging authors to cover publishing costs (currently US\$2750 and rising). PLoS are also now receiving extra finance from advertisements and donations. Once again, the cost issue has re-surfaced and is forcing OA journals to compromise their initial ideological stances. Although they may be 'not-for-profit', PLoS is certainly neither cost free, nor cheap.

A total of 182 biomedical OA journals are included in BioMed Central (www.biomedcentral.com/). Importantly, the golden road requires the consent of the participating peer-reviewing journals, who usually retain the copyright (distributed under a Creative Commons or similar license that generally allows more liberal use than is permitted for a traditional copyrighted work). Furthermore, the release time of published articles can vary, since some of the parti-

icipating journals impose embargoes whereby articles are only freely accessible after a set period of time (that can range from 4 to 36 mo). The golden road may be free for readers, but increasingly it is becoming more and more costly for authors.

One may argue that editors of OA golden road journals, pressured by their commercial employers, may be forced to accept articles of lower quality in order to increase the number published and therefore the journal's income. However, it is clear that this is double-edged, since such a tactic can result in a drop in JIF (which would make the journal less attractive). The golden road is arguably discriminative towards authors from developing countries. This situation is amplified for the top journals, since their high rates of rejection mean that the small number of publishing authors need to bear the cost.

A 'platinum road' has also been proposed, whereby OA publishing is subsidised totally through author fees. However, we see this as simply one cost model of the golden road. The costing of the 'not-for profit' PLoS is an example. Of the 307 publishers registered at the OA monitor—ROMEIO (<http://romeio.eprints.org/stats.php>)—63% (representing 6381 journals) are 'green', providing OA to readers for pre- and post-prints without author costs. A further 29% (2909 journals) are 'pale green', meaning that they provide OA only for readers of pre-prints. Although these statistics are impressive, we wish to stress that the debate about OA should not only be based on the provision of free access to articles for readers, but that it should also promote free economic access to authors to publish. At present, only the green road based on self-archiving is cost-free to authors and readers (see also section 'The economics of open access') and is therefore the priority for the 'OA movement'. Furthermore, there are already at least 10 different service providers like SCOPUS using the OAI-PMH standard (Van de Sompel & Lagoze 2000). This means that their archives (typically self-archived institutional output or journal/publisher databases) can, and are, being inter-linked.

From its utopian origins, OA is now a contending publishing practice. One key reason for its popularity is its 'citation advantage' described in the next section.

Citation advantage

Harnad (2001a) asserts that:

'Researchers publish their findings in order to make an impact on research, not in order to sell their words ... to make a difference, to build upon the work of others, and to be built upon in turn by others. 'So, it is in the interest of the author to have the greatest exposure, or impact, for the literature they produce.'

It is not the number of articles published that reflects the return on research investment. A piece of research, if it is worth funding and doing at all, must not only be published, but used, applied and built upon by other researchers, worldwide. This is called 'research impact' and one measure of it is the number of times an article is cited by other articles, i.e. 'citation impact' (Harnad 2006b). Citation advantage is the notion that OA increases citation impact.

The first research into whether or not OA articles have higher citation impact was published in 2001. Lawrence (2001) analysed the difference in citation impact between articles freely available to readers on the web and those only available through either toll-access services or paid-for paper print-form. Lawrence found that the OA articles received on average 2.6 times more citations. In addition, Kurtz has shown that in astrophysics, a field in which there is already effectively 100% OA (due to support from institutional licensing), the overall 'usage' (citations and downloads) of papers doubled following the move to OA (Kurtz 2004). This finding was soon extended beyond physics and computer science (Harnad et al. 2004) to 10 more subject areas, including biological and social science, and humanities (Hajjem et al. 2005).

The evidence so far shows that an increase in citation of OA material due to greater visibility and accessibility is resulting in a citation advantage. Whether the citation advantage (currently 260%) is a short-term transition effect or not remains to be seen. However, as we explain in the section 'Evaluating research performance using OA', new OA modes of publishing and associated scientometric indices are expected to maintain the citation advantage.

The dollar value of a citation

In today's very highly competitive academic environment, it is clear that not only do you have to publish in journals with high JIFs to get funding, but your articles also need to be highly cited to have an impact. Citations then, clearly translate into cash. In 1986, before the big wave of self-archiving and the adoption of the green road to OA, it was shown that the marginal dollar value of a citation to articles printed by commercial journals in the USA was estimated to be between US\$ 50 and 1300 depending on the field (Diamond 1986). Scaling this up by 170% to account for inflation in the period 1986 to 2005, then the dollar value of a citation has risen to approximately US\$ 90 to 2200. Although these figures may be surprising in themselves, their cumulative effect when taken in the context of world-wide or national research is astounding. For example, the UK research councils

spend £3.5 billion annually, which results in an average of 130 000 journal articles being published per year (ISI figures) with an average citation rate of 5.6. This corresponds to 761 000 citations. Self-archiving increases citation impact by 50 to 250% but, so far, only 15% of researchers are spontaneously self-archiving their articles (Harnad et al. 2004). Taking the lower estimate, we find that the loss of potential research impact due to the 85% of authors not self-archiving in the UK is $50 \times 80\% \times \text{£}3.5 \text{ billion} = \text{£}1.5 \text{ billion}$. Hence, it is possible to argue that the green road to OA is a source of wealth creation.

The economics of OA

Exactly what are the real costs associated with the green road? Harnad (2001a) describes a system or network of public institutional e-print archives (with authors depositing their work in their local archive) to be harvested by federating services (such as OAI-PMH) into what he calls 'global virtual archives'. These archives are toll-free (to the reader) accessible web repositories of literature deposited by their own authors.

However, costing models for such repositories are still embryonic. For example, Fried-Foster & Gibbons (2005) cite MIT (Michigan Institute of Technology) libraries as estimating an annual maintenance cost for their institutional repository of US\$285 000, but with 'approximately 4000 items currently housed, this translates to over US\$71 spent per item, per year.' The cost of establishing an institutional repository varies greatly across implementations. At the free end, open source software is available to allow authors to self-archive on personal webpages. For example, Carr & Harnad (2005) found that authors from a Computer Science department were taking just 10 min to deposit each paper. Another good sign is that the monthly submission rate of articles archived on arXiv has now reached 5000 (arXiv, <http://arxiv.org>).

Free software is also available for producing an institutional repository, an OA journal or even a complete digital library. The problem is that although the software is free, the labour, associated management and storage costs paid by host institutions, mean that the green road can never be entirely cost-free. In the section 'In defence of public science', we propose a new publishing model that, by redistributing costs, could solve the costing crisis that the green OA movement is facing.

If OA can be shown to maximise research impact, then there is a strong empirical and economic argument to convince authors and institutions of the benefit of providing OA to their work. Even the UK and US

governments have recognised the benefit of increased access to research papers, in particular that tax-payers should not have to pay to access the results of research that they fund in the first place (Suber 2006).

Given that an economically viable and sustainable solution for the provision of OA is possible (see the section 'In defence of public science'), what is the current status of OA-based bibliometrics?

Evaluating research performance using OA

New 'scientometric' indicators of digital impact are emerging to chart the online course of knowledge. For example, the OpCit exercise, funded by the Joint NSF-JISC International Digital Libraries Research Programme, was a 3 yr project that began in October 1999 and was completed at the end of 2002 (Hitchcock et al. 2002). The project examined and developed 2 key tools for OA bibliometrics: Citebase and GNU EPrints.

Citebase

Citebase is a citation-ranked search and impact discovery service that measures citations of scholarly research papers that are available on the internet in multi-disciplinary archives such as arXiv, CogPrints and BioMed Central. It harvests OAI-PMH metadata records for papers in these archives and automatically extracts the references from each paper. The combination of data from an OAI-PMH record for a selected paper with the references from and citations to that paper is the basis of the Citebase record for the paper. Citebase provides 4 services: a graph of the article's citation/hit history, a reference list of all articles cited by each article, the top 5 articles citing an article, and the top 5 articles co-cited with an article. The association between document records and references provides the basis for a classical citation database of the future. Citebase has records for 230 000 papers, indexing 5.6 million references. By discipline, approximately 200 000 of these papers are classified within arXiv physics archives (Hitchcock et al. 2002).

GNU EPrints

GNU EPrints is software for building OA archives and is aimed at institutions and special-interest communities. It is currently used by nearly 60 different archives (Hitchcock et al. 2002). It is open source and freely available under the GNU General Public License. GNU EPrints provides internationalised metadata stored as Unicode to open up access to less-widely

recognised publishing languages. It also provides support for housing multiple archives on 1 server. Hitchcock et al. (2002) state:

'EPrints is extending its focus on institutional research papers. It is now configurable for adoption as a journal-archive, e.g. Behavioral and Brain Sciences and Psychology, by new open access journals or established journals converting to open access, and will include the facility to manage peer review and peer commentary'

In addition to these tools from the OpCit Project, autonomous and free OA citation tools, such as Cite-seer (Lawrence et al. 1999), mean that information scientists can now build comprehensive citation databases, limited only by what is already OA. The potential benefit that OA has for service provision has become clear. As authors self-archive or publish in OA journals, so the value and potential of services built on that literature grows. Services built on OA therefore compete not on the basis of how much content they contain, but on the quality of the service that they provide, free of licensing barriers. This leads to a user-driven environment where the real value of citations becomes transparent for all.

At the more general level, OA could provide 3 benefits (above and beyond the current system) for research evaluation. Firstly, the inter-linking of institutional repositories can lead to the creation of a unified platform for administering the collection and submission of research papers to be evaluated (Carr & MacColl 2005). Secondly, evaluation methods that include citations would become easier to perform due to the deeper search capabilities of engines like Cite-seer. Importantly, this means that authors can be informed about their citation impact without having to pay subscriptions to ISI for this important and personal information. Thirdly, by making all of the information used for evaluation (the full-text papers, peer review comments and revisions) transparent, OA may help to create a visible and accessible and fair nexus for the evaluation of research.

Economic restrictions, the desire for 'fair play', the potential offered by OA for citation advantage together with free software, have created the critical mass that have led to an exodus by academics from commercial publishers.

Declaring independence

A number of organizations such as Create Change, the Scholarly Publishing and Academic Resources Coalition (SPARC), and the BOAI are all openly pushing for OA. In an important document published by SPARC, entitled, *Declaring Independence: A guide to creating community-controlled journals* (SPARC 2000),

Michael Rosenzweig, the Editor-in-Chief of the journal *Evolutionary Ecology*, issued the following call to action:

'We scientists can exercise control of our journals. We can transform them from commercial commodities back to instruments of service to education and research. When we are in control, we fulfill our responsibility to ourselves, to society, to our institutions, and to our colleagues throughout the world... We scientists are naturally curious explorers. Please be curious about the way your journal is run. And if you don't like what you find, explore ways of declaring independence'

The green road being followed by authors, the setting up of OA journals and repositories, the development of OA citation software, as we have seen, comprise the empirical evidence of the exodus from commercial publishers and the practical end of the declaration of independence. OA is ushering in a new epoch for scholarly publishing where finance has become the new arena in the battle between scientists wanting to pay less and commercial journals wanting to profit more.

JOURNALS GO ON THE OFFENSIVE

The declaration of independence by many authors and journal users represents a loss of revenue for the journals. Moreover, as we have seen, the new OA initiatives are gaining support and numbers. In the face of profit losses, journals are hitting back. We reveal some of the tactics that journals are adopting in an attempt to defeat the OA movement.

The 'open choice' Trojan Horse

What is wrong with publishers lobbying the US Congress to stop the federal government from providing OA to publicly funded research? Lessig (2006) in a column to *Wired* puts it this way:

'Imagine if tyre manufacturers lobbied against filling potholes so they could sell more tyres... or what if taxicab companies managed to get a rail line placed just far enough from an airport to make public transportation prohibitively inconvenient?'

After years of lobbying against OA and failing, journals have sent in a Trojan Horse. It is called the 'open choice' option, i.e. the option for authors to buy OA for their individual articles. For example, the journal PNAS (*Proceedings of the National Academy of Sciences*) on their website (www.pnas.org) state:

'Authors of research articles may pay a surcharge of \$1,100 to make their paper freely available through the PNAS open access option... The surcharge is intended to cover administrative costs and potential lost revenue.'

Not surprisingly, with large profits to be gained (US\$ 1100 per article), there has been a spate of publishers announcing that they are offering the open choice or 'OA option'. However, this should be seen in the global context of OA self-archiving mandates that are pending in the US, UK, Australia, Germany, France and Norway. Open choice is a Trojan Horse; it is the publishers' way of lobbying for paid-OA rather than self-archiving. It is a clever, but somewhat cynical way of delaying still longer the immediate mandating of OA (Harnad 2006a). The current popularity of the OA option is ironic. In the competitive world of scholarly publishing, authors seem willing to pay for OA in exchange for the return it gives them in terms of citations and subsequent career opportunities. Far from resolving the problem of the competition they face, paid-OA continues to reinforce the journal hierarchy with costs simply being re-distributed from the user to the author. The outcome, we believe, will be the same—preferential attachment to high JIF journals (now OA) and potentially stronger competition. Journals are shifting the economic costs to authors and attempting to co-opt the golden road. As we shall see in the section 'In defence of public science', there could be alternative cost models for publishing, where authors and readers do not have to pay and where the management costs can be met.

Proprietary preprint servers

To fight against the green road and the popular article repositories such as arXiv and Citeseer, giants like Reed Elsevier have tried setting up their own competing preprint archives. Examples included 'The Chemistry Preprint Server' (now 'ChemWeb', run by Chem-Industry.com), the 'Computer Science Preprint Server' and the 'Mathematics Preprint Server'. The latter was especially devious, because mathematicians refer to arXiv as 'the mathematics preprint server'. However, the Elsevier preprint servers did not convince mathematicians, chemists and computer scientists, who boycotted them and continued submitting to arXiv and Citeseer (as described in sections 'Open access (OA)' and 'The economics of open access').

Eventually Elsevier gave up and announced (at: www.sciencedirect.com/preprintarchive):

'Despite their wide readership, the Chemistry, Maths and Computer Science research communities did not contribute articles or online comments to the Preprint service in sufficient numbers to justify further development. Consequently on the 24th of May, 2004 the 3 Elsevier Preprint Servers - Chemistry, Math and Computer Science - stopped accepting new submissions to their sites. The current site is now a freely available and permanent web archive for those research articles already submitted to the Preprint Servers'

Bundled library journal subscriptions

Reed Elsevier's principle method of selling journal subscriptions to libraries is to sell them as huge all-or-nothing packages like Science Direct that contains as they say, 'over 2000 journals' (www.sciencedirect.com/). However, as we have already seen, with some 24 000 journals in circulation (Harnad et al. 2004), these bundles represent a tiny minority of all journals. The bundles mean that libraries are denied the option of saving money by picking and choosing those journal subscriptions demanded by their users. As a result, in 2003 Cornell University cancelled their subscription to all Elsevier journals, and 4 North Carolina universities have also joined the boycott (Suber 2006).

Using the law

Gordon and Breach filed a lawsuit initiated in 1988 against the American Physical Society (APS) for publishing lists of journal prices. Under the slogan of 'competitive disadvantage', the publishers went to court to try to prevent their financial data from being made publicly available. The case lasted 9 yr, but finally APS won. As a result of this victory, scholars can now inform themselves of journal prices and choose not to work for those journals they consider to be over-priced.

Another landmark case is the case of Eric Weisstein of Mathworld (Weisstein 2001). On March 8, 2000, Weisstein was served with a document naming him and his employer (Wolfram Research) as defendants in a Federal copyright-violation lawsuit with his publisher (CRC Press LLC, formerly the Chemical Rubber Company) after signing a contract to publish *The Concise Encyclopaedia of Mathematics*—a 'snapshot' in book form of his website mathworld.wolfram.com. In its lawsuit, CRC claimed that the existence of the MathWorld website, 'competes with and interferes and impairs with sales of the Concise Encyclopedia' (Weisstein 2001). Remarkably, Weisstein lost the case, even though his website preceded the book. He had unwittingly signed away all rights to the publisher. The point is that whatever rights you do not explicitly keep as an author, the publisher will get. Authors need to be on guard and highly aware in the world of for-profit journals.

Propaganda

Besides the anti-OA lobbying strategies, large journal publishing companies are now resorting to pure propaganda. Recently, the American Association of Publishers (AAP) hired pro-business public relations expert Eric Dezenhall. According to Giles (2007):

'Eric Dezenhall has made a name for himself helping companies and celebrities protect their reputations, working for example with Jeffrey Skilling, the former Enron chief now serving a 24-year jail term for fraud.'

However, Weiss (2007) shows that they are very much on the defensive:

'...there is a potential downside to hiring the likes of Dezenhall: If word gets out, you stand to be seen as on the ropes and willing to do anything to win.'

The situation is very serious. In the profit-ruled world, science faculties are being forced to close, financially firewalled libraries are cutting back on journal subscriptions, authors are being pushed to pay publishing fees and OA-option costs are spiralling. In the OA world, despite the development of free software like Citeseer and GNU Eprints, the rate of self-archiving is still low and viable costing models of the green road are still to be addressed. Worse still, the outcome of the battle between the commercial journals and the OA movement is likely to determine the future of academia, i.e. whether or not it will become the sole preserve of a financial elite who can afford to perform research and communicate it, rather than a public arena where scholarly ideas can be created, produced and freely accessed.

In the next section, we describe ideas and strategies that have been put forward in an attempt to guarantee the success of the OA movement and to ensure future public access to the results of research.

IN DEFENCE OF PUBLIC SCIENCE

A blueprint for a 'green' OA publishing model

As the focus of the author self-archiving OA strategy has moved from centralised, subject-specific archives towards developing institutionally hosted archives, so it has become increasingly useful to monitor the progress of these disparate resources. The Registry of Open Access Repositories (<http://roar.eprints.org/>) provides a compilation of such archives, and Celestial (<http://celestial.eprints.org/status>), designed to harvest metadata from source repositories using OAI-PMH software, monitors the number of records available. Combine this with the availability of open-source journal publishing software like GNU Eprints, then the time is ripe to start discussing new publishing models.

To stimulate discussion over such future directions, we have sketched a blue-print for such a model. This is presented in Fig. 3.

The key to this model is that all costs derive from the research funding offsets that result from the absence of the need to pay for journal subscriptions, author publishing charges (including the 'open access option') and local archival costs. Although free open-source journal publishing software is available, as we have described, the labour involved in developing international journals with it of course has associated costs (one might also include translation software and imple-

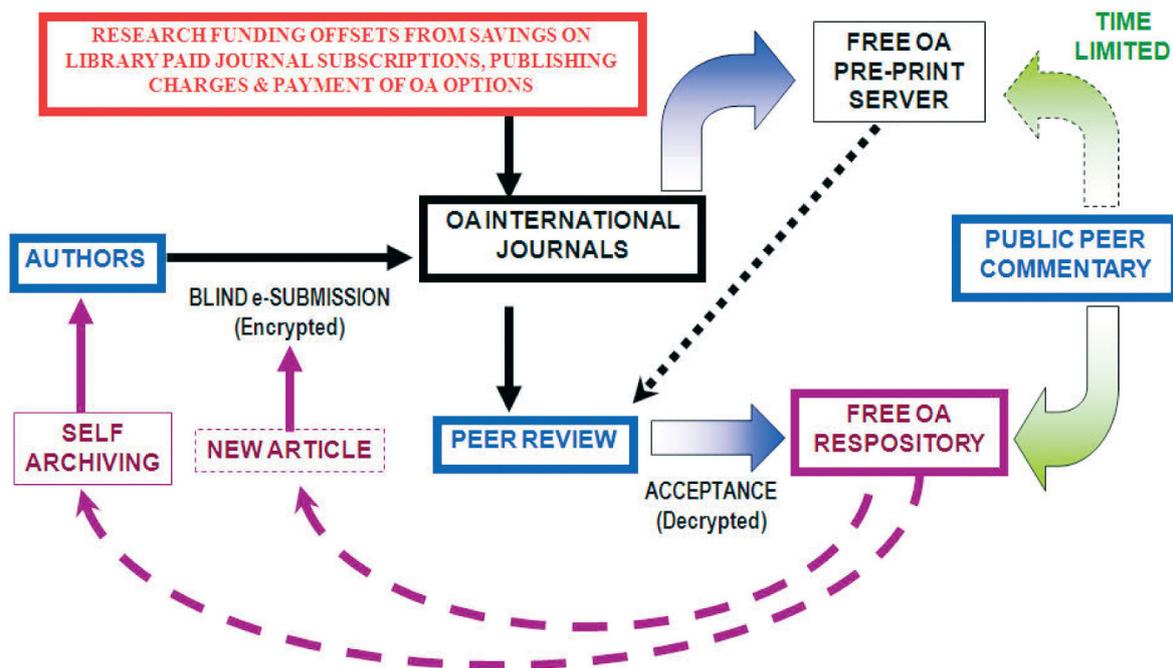


Fig. 3. A blueprint model for autonomous and open access (OA) academic publishing

mentation costs, copyediting and typesetting costs and the cost of managing the repository and paying peer-reviewers, for example). Of course such work does not need to be voluntary or pro-bono, as the savings from research funding offsets can meet these economic needs. Data storage is a material cost that also needs covering. Assuming that research funding offsets are capable of covering such costs, the publishing cycle then proceeds as depicted in Fig. 3. We have attempted to construct the model, bearing in mind the issues we have raised considering the interplay between authors, editors and reviewers mentioned in the section 'Biomedicine: at the rough end of the journal monopoly'.

Authors submit 'blind' their articles using a LaTeX-like template file that encrypts the author names and affiliations and which produces a decryption key kept by the author. Submitted articles are then sent out for peer review to members of public and professional organisations in the subject area of the electronic journal (perhaps selected by an elected committee) by employees of the OA repositories. The article (with encrypted authorship) is made freely available on a subject-based repository preprint server where discussion threads and commentary from readers may be posted during a limited timeframe (perhaps 1 to 2 mo). One could argue that in small fields where authors are well known to editors, real blind submission may not be realistic. Still, new researchers or research groups will benefit from this system. The peer reviewers, taking also into account the public peer commentary, will decide whether or not to accept the article for inclusion in the OA repository of the electronic journal, at which point the authorship is decrypted on provision by the authors of their key. Public peer commentary, although time-consuming, is of great importance for scientific debate and communication. Furthermore, it is already in existence in a small number of journals and even *Nature* has flirted with the idea. Its lack of widespread adoption is most likely due to scientists' traditional way of perceiving scientific communication through closed peer review and the belief that there is no other way. But this does not invalidate the concept or its potential benefits. Peer commentary may continue on the article in the electronic journal repository as and if desired and may stimulate the authors along new directions of investigation, leading possibly to submission of a new article that has benefited from the openness inherent in the model. The actual nature of the discussion will develop in step with new developments of web dialogue. For example, blogs, discussion threads and forums are commonplace on the internet and their software can be integrated into the publishing process and managed by the OA repositories.

We welcome suggestions and constructive criticisms from readers so that a viable and sustainable (green) OA publishing model can be developed and implemented.

A glimpse of the future

All refereed journals will soon be available online; most of them already are. This means that potentially anyone will be able to access them from any networked computer. The OA literature will soon be interconnected by citation (with OpCit link counters), author and keyword/subject links, allowing for unheard-of power and ease of access and navigability. Successive drafts of preprints will be linked to the official refereed version as well as to any subsequent updates enhancing the self-correctiveness of scholarly and scientific research in remarkable new ways. BOAI-compliant free software is now available and institutions can immediately create Eprint archives in which their authors can self-archive all their refereed papers. Harnad (2001b) states:

'These archives (www.openarchives.org) will then be harvested into global, jointly searchable 'virtual archives' that are likely to be more interactive and productive, spawning their own new digital metrics of productivity and impact.'

Just as there is no longer any need for research or researchers to be constrained by the access-blocking restrictions of paper distribution, there is no longer any need for authors and readers to be constrained by paid access either. OA is achievable and, most importantly, it is already partly implemented. The OA ideal would provide entire full-text refereed publications online without financial access barriers (to readers and authors) with all articles citation-interlinked and fully searchable and retrievable.

We have proposed a publishing model that may accomplish this with a practical solution for the costs involved. At the institutional level, it can be implemented immediately so as to resolve the 'serials crisis' (see section 'The crisis'). The dynamics of the publishing process (submission and peer review) can also be worked out between federating repositories and OA journals; we look forward to such initiatives.

In the meantime, there are steps that the individual scholar can take immediately (inspired by Harnad 2001c): (1) choose not to publish in journals you consider to be over-priced (using perhaps journalprices.com as a guide); (2) choose not to do free work for commercial journals such as editing and refereeing unless they pay for your overheads, as mentioned in the section 'The birth of autonomous scientific publishing'. Better still, perform these tasks for green OA journals; (3)

choose to participate only with those journals that allow you (the author) to self-archive and to place your paper in an OA repository, such as arXiv; (4) choose to use open-source software and help to set up OA journals where they are needed; (5) pressure your institution to install and register OAI-PMH compliant e-print archives; and (6) adopt the preprint-corrigenda strategy to avoid paying author costs for peer-reviewed articles.

These steps are simple to implement and help ensure that (1) academics can publish their publicly funded work without paying fees and (2) such work will be freely available to anyone who wishes to read it.

CONCLUSION

The current system of commercial academic publishing, as we have shown, is seriously flawed. Alternatives, such as the author-pays OA golden or platinum roads, have also been shown to be problematic and indeed may accentuate the journal hierarchy. The publishing model we propose is born out of the needs of authors and readers faced with these problems. Whether or not proposals of this nature will be adopted by the scientific community, we believe, is a question of how much higher publishing barriers will be in the future.

At some time or other, everyone involved in the publishing process—from the reader to the head of academic institutions—has faced the problem of locating an article from a journal that prohibits access. Ask any researcher about the limitations involved when submitting a paper, getting published in high JIF journals and succeeding in grant applications or securing tenure and a pattern emerges: one of increasing difficulty. However, the OA movement, as we have shown, is not without problems of its own. Self-archival rates are still low, and authors are continuing to pay for OA options and high publication charges from OA journals. Perhaps it is lack of awareness or more importantly a symptom of the competitive pressures that authors face in a world where their h-index and personal JIF rules.

We have presented the facts that reveal the magnitude of the crisis resulting from the commercialisation of science as well as some of its consequences—the overpricing of academic journals, financial-firewalling and a systematic tendency towards ‘blind’ evaluation by numbers, such as the JIF and the h-index. We have demonstrated how the present evaluation system based on citation indices is converting all scientific output into a quantifiable product, owned and provided at high cost by profit-making

organisations such as ISI. Citation impact factors are being translated into money, prestige and job positions, and scholars are being pressured into investing their creativity and effort in the publishing game rather than science.

To combat the hegemony of the JIF and the h-index, new multi-parameter bibliometrics are being proposed but, as we have argued, until a universal definition of quality is agreed on this will remain problematic. Bibliometrics still seem to have a place but they need to evolve and adapt in order to match the complexity and diversity of academic publishing in the new digital age. The new interpretations from social network theory and multi-parameter studies for example are testament to this. That power laws and hierarchies exist, in themselves, is not surprising. After all, scientific ideas evolve, undergo natural selection, at times fit the environment of academic consensus and at other times revolutionise it. The same argument can apply to scientific journals, but without a measure of quality it is simply commerce that rules. It is this distaste that has, in part, led to dissent amongst the ranks of academics.

As with all conflicts, change is inevitable. New trends have emerged on the academic publishing landscape—self-archiving, OA journals, OA bibliometric indices, preprint servers and new publishing models where authors are in the driving seat. Scholars are becoming aware not only of the schism between the 2 worlds of the commercial enterprises and the scientific community, but also of the sustainable opportunities to be financially independent. Highly esteemed academics along with young emerging scientists are dedicating much of their precious academic time to bibliometric studies in an effort to raise the consciousness of their peers about the situation.

Although the majority of scholars still feel powerless in the face of the journal monopoly, small and important battles have already been won. We have given some examples in our review of some of the creative fruits born from the needs of scholars, such as the green road and free software for setting up e-journals and harvesting metadata from them; we will no doubt see more initiatives like these in the future.

For now at least, the siege of science continues. We have speculated how it may evolve and develop and why the single-parameter indices of old may cease to exist. However, it is clear is that the OA movement is gathering under a common banner and that its message is uniform: as Tuck (2003) states:

‘There seems to be no appreciation of the once unchallenged argument that the scientific goose that lays the golden eggs needs some tall green grass, privacy and free choice of nesting sites’

Acknowledgements. This article would not be what it is without the hard work and commitment of Howard Browman and Konstantinos Stergiou, to whom we owe a great debt of thanks. We believe that a critical analysis of bibliometrics was much needed and we feel very proud to appear in this Theme Section. We also thank the European Association of Science Editors (EASE) for their financial support and, in particular, Catherine Mark of the CNB-CSIC for her enthusiasm and for all that we have learnt from her expertise in translating scientific ramblings into plain English. Finally, as in all scientific works, we thank all those who have provided the large body of research we could draw upon and those who, often at risk to their own academic careers, exposed the dangers to knowledge posed by bibliometric abuses, innumeracy and corruption, as well as those who helped to develop practical solutions and alternatives—in particular Peter Lawrence and Stevan Harnad, whose articles provided the initial motivation for our own research.

LITERATURE CITED

- Agres T (2005) When the line between science and business blurs. With conflicts of interest making headlines, can trust be restored? *Scientist* 19:42–47
- Barabasi AL, Albert R (1999) Emergence of scaling in random networks. *Science* 286:509–512
- Barcinski MA (2003) Disruption to science in developing countries. *Nature* 423:480
- Bauer HH (2004) Science in the 21st century: knowledge monopolies and research cartels. *J Sci Explor* 18: 643–660
- Bergstrom CT (2007) Eigenfactor: measuring the value and prestige of scholarly journals. *College and Research Libraries News* 68(5)
- Bergstrom CT, Bergstrom TC (2004) The costs and benefits of site licences to academic journals. *Proc Natl Acad Sci USA* 101:897–902
- Bergstrom TC (2001) Free labor for costly journals? *J Econ Perspect* 15:169–182
- Bergstrom T, McAfee RP (1997) An open letter to all university presidents and provosts concerning increasingly expensive journals. Available online at: www.hss.caltech.edu/~mcafee/Journal/OpenLetter.pdf
- Bollen J, Van de Sompel H, Smith JA, Luce R (2005) Toward alternative metrics of journal impact: a comparison of download and citation data. *Inf Process Manag* 41: 1419–1440
- Bollen J, Rodriguez MA, Van de Sompel H (2006) Journal status. *Scientometrics* 69:669–687
- Bradford SC (1934) Sources of information on specific subjects. *Engineering—an illustrated weekly journal (London)* 137:85–86
- Brody TD (2006) Evaluating research impact through open access to scholarly communication. PhD thesis, University of Southampton
- Brookfield J (2003) The system rewards a dishonest approach. *Nature* 422:259–261 [PubMed doi:10.1038/422259a](https://doi.org/10.1038/422259a)
- Buela-Casal G, Perakakis P, Taylor M, Checa P (2006) Measuring internationality: reflections and perspectives on academic journals. *Scientometrics* 67:45–65.
- CAAT (Campaign Against the Arms Trade) (2007) Campaign success: Reed Elsevier to stop organising arms fairs. Available online at www.caat.org.uk/campaigns/reedelsevier.php
- Calza L, Garbisa S (1995) Italian professorships. *Nature* 374:492
- Campanario JM (2003) Coping with Peer Rejection. *Nature* 425:645. Also available online at: <http://www2.uah.es/jmc/nobel.html>
- Campbell P (2008) Escape from the impact factor. *Ethics Sci Environ Polit* 8: (in press)
- Carr L, Harnad S (2005) Keystroke economy: a study of the time and effort involved in self-archiving—technical report. University of Southampton. Available online at: <http://eprints.ecs.soton.ac.uk/10688/>
- Carr L, MacColl J (2005) IRRA (Institutional Repositories and Research Assessment) RAE software for institutional repositories—technical report. IRRA Project. Available online at: <http://irra.eprints.org/white/>
- Chan L, Kirsop B (2001) Open archiving opportunities for developing countries: towards equitable distribution of global knowledge. *Ariadne* 30. Available online at: www.ariadne.ac.uk/issue30/oai-chan/
- Chan L, Cuplinskas D, Eisen M, Friend F, and others (2002) Budapest Open Access Initiative. Available online at: www.soros.org/openaccess/read.shtml
- Dalton R (2001) Peers under pressure. *Nature* 413:102–104
- Darmoni SJ, Roussel F, Benichou J, Thirion B, Pinhas N (2002) Reading factor: a new bibliometric criterion for managing digital libraries. *J Med Libr Assoc* 90:323–327
- Delamothe T, Smith R (2001) PubMedCentral: creating an Aladdin's cave of ideas. *BMJ* 322:1–2
- Diamond AM Jr (1986) What is a citation worth? *J Human Resour* 21:200–215
- East JW (2006) Ranking journals in the humanities: an Australian case study. *AARL* 37(1):3–16
- Egghe L (2006) Theory and practice of the g-index. *Scientometrics* 69:131–152
- Faloutsos M, Faloutsos P, Faloutsos C (1999) On power-law relationships of the internet topology. *SIGCOMM* 29: 251–262
- Fassoulaki A, Paraskeva A, Papilas K, Karabinis G (2000) Self-citations in six anaesthesia journals and their significance in determining the impact factor. *Br J Anaesth* 84:266–269
- Feder G, Rohde JE, Sebastián MS, Janlert U and others (2005) Reed Elsevier and the international arms trade. *Lancet* 366:889
- Fried-Foster N, Gibbons S (2005) Understanding faculty to improve content recruitment for institutional repositories. *D-Lib Magazine*, 11(1). Available online at: www.dlib.org/dlib/january05/foster/01foster.html
- Garfield E (1955) Citation indexes for science: a new dimension in documentation through association of ideas. *Science* 122:108–111
- Garfield E (1972) Citation analysis as a tool in journal evaluation. *Science* 178:471–479
- Garfield E (1979) *Citation indexing*. John Wiley & Sons, New York
- Garfield E (1990) How ISI selects journals for coverage: quantitative and qualitative considerations. *Current Contents*, 28 May. Available online at: www.garfield.library.upenn.edu/essays/v13p185y1990.pdf
- Garfield E (1998) Der Impact Faktor und seine richtige Anwendung. *Unfallchirurg* 101:413–414
- Garfield E (2005) The agony and the ecstasy—the history and the meaning of the journal impact factor. *Int Congr Peer Review Biomedical Publ*, Chicago, IL, Sep 16. Available online at: www.garfield.library.upenn.edu/papers/jifchicago2005.pdf
- Garfield E, Sher IH (1963) New factors in the evaluation of scientific literature through citation indexing. *American Documentation* 14(3):195–201. Also available online at: www.garfield.library.upenn.edu/essays/v6p492y1983.pdf

- Giles J (2007) PR's 'pit bull' takes on open access—journal publishers lock horns with free-information movement. *Nature* 445:347. Available online at: www.nature.com/nature/journal/v445/n7126/full/445347a.html
- Gross PLK, Gross EM (1927) College libraries and chemical education. *Science* 66(1713):385–389
- Guédon JC (2002) In Oldenburg's long shadow: librarians, research scientists, publishers, and the control of scientific publishing. Available online at: www.arl.org/resources/pubs/mmproceedings/138guedon.shtml
- Hajjem C, Harnad S, Gingras Y (2005) Ten-year cross-disciplinary comparison of the growth of open access and how it increases research citation impact. *IEEE Data Engineering Bull* 28(4):39–47. Available online at: <http://eprints.ecs.soton.ac.uk/11688/>
- Hall P (1998) *The innovative milieu, in cities in civilization*. Parthenon, New York
- Hansen HF, Jorgensen BH (1995) *Styring af forskning: kan forskningsindikatorer anvendes?* Samfundslitteratur, Copenhagen
- Harnad S (1995) Universal FTP archives for esoteric science and scholarship: a subversive proposal. In: Okerson A, O'Donnell J (eds) *Scholarly journals at the crossroads; a subversive proposal for electronic publishing*. Association of Research Libraries, Washington, DC
- Harnad S (2001a) The self-archiving initiative: freeing the refereed research literature online. *Nature* 410:1024–1025. Available online at: www.nature.com/nature/debates/e-access/Articles/harnad.html
- Harnad S (2001b) For whom the gate tolls? How and why to free the refereed research literature online through author/institution self-archiving, now. Available online at: <http://cogprints.org/1639/>
- Harnad S (2001c) Minotaur: six proposals for freeing the refereed literature online: a comparison. *Ariadne* 28. Available online at: <http://www.ariadne.ac.uk/issue28/minotaur/>
- Harnad S (2003) Open access to peer-reviewed research through author/institution self-archiving: maximizing research impact by maximizing online access. In: Law D, Andrews J (eds) *Digital libraries: policy planning and practice*. Ashgate Publishing, Southampton
- Harnad S (2006a) Open choice is a Trojan Horse for open access mandates. *SPARC Open Access Forum*, Jun 28. Available online at: <http://users.ecs.soton.ac.uk/harnad/Hypertext/Amsci/5443.html>
- Harnad S (2006b) Publish or perish—self-archive to flourish: the green route to open access. *ERICIM News* 64
- Harnad S (2007) Open access scientometrics and the UK Res 11th Annu Meet Int Soc Scientometrics and Informetrics, Madrid. Available online at: <http://arxiv.org/ftp/cs/papers/0703/0703131.pdf>
- Harnad S, Brody T, Vallieres F, Carr L and others (2004) The access/impact problem and the green and gold roads to open access. *Serials review* 30(4). Available online at <http://eprints.ecs.soton.ac.uk/10209/>
- Hecht F, Hecht B, Sandberg AA (1998) The journal 'Impact Factor': a misnamed, misleading, misused measure. *Cancer Genet Cytogenet* 104:77–81
- Hirsch JE (2005) An index to quantify an individual's scientific research output. *Proc Natl Acad Sci USA* 102(46):16569–16572
- Hitchcock S, Bergmark D, Brody T, Gutteridge C, and others (2002) Open citation linking: the way forward. *D-Lib* 8(10). Available online at: www.dlib.org/dlib/october02/hitchcock/10hitchcock.html
- Hofbauer R, Frass M, Gmeiner B, Kaye AD (2002) Euro-Factor (EF)™—the European Journal Quality Factor: the new European 'scientific currency'. *VICER Publishing, Vienna*
- Horton R (2000) The controversy over genetically modified food exposes larger issues about public trust in science and the role of science in policymaking. *Med J Aust* 172:148–149
- Insall R (2003) Impact factors: target the funding bodies. *Nature* 423:585
- Jennings C (1998) Citation data: the wrong impact? *Nat Neurosci* 1:641–642
- Jeong H, Mason SP, Barabasi AL, Oltvai ZN (2001) Lethality and centrality in protein networks. *Nature* 411:41–42
- Jiménez-Contreras E, Delgado E, López-Cózar E, Ruiz-Pérez R, Fernández VM (2002) Impact-factor rewards affect Spanish research. *Nature* 417:898. Available online at: <http://dx.doi.org/10.1038/417898b>
- Kelly CD, Jennions MD (2006) The h-index and career assessment by numbers. *Trends Ecol Evol* 21:167–170
- Kirby R (1997) Letter to Elsevier officers. *Newsletter on Serials Pricing Issues* 199:1
- Kurtz MJ (2004) Restrictive access policies cut readership of electronic research journal articles by a factor of two. Available online at: <http://opcit.eprints.org/feb19oa/kurtz.pdf>
- Kyrillidou M, Young M (2004) *ARL Statistics 2003–04*. Tech rep 2004. Available online at: http://www.arl.org/sparc/bm~doc/Cooperatives_v1.pdf
- Laherrere J, Sornette D (1998) Stretched exponential distributions in nature and economy: 'Fat tails' with characteristic scales. *Eur Phys JB* 2:525–539
- Lavalette D (1996) *Facteur d'impact: impartialité ou impuissance?* Internal Report, INSERM U350, Institut Curie, Paris
- Lawrence PA (2003) The politics of publication. Authors, reviewers and editors must act to protect the quality of research. *Nature* 422:259–261
- Lawrence S (2001) Online or invisible? *Nature* 411:521
- Lawrence S, Giles CL, Bollacker K (1999) Digital libraries and autonomous citation indexing. *IEEE Computer* 32:67–71
- Lehmann S, Jackson AD, Laurup B (2005) Life, death and preferential attachment. *Europhys Lett* 69:298–303
- Lessig L (2006) *Crushing Competition*. *Wired Issue* 14:05
- Liljeros F, Edling CR, Nunes-Amaral LA, Stanley HE, Åberg Y (2001) The web of human sexual contacts. *Nature* 411:907–908
- Linardi PM, Coelho PM, Costa HM (1996) The 'impact factor' as a criterion for the quality of the scientific production is a relative, not absolute, measure. *Braz J Med Biol Res* 29:555–561
- Lipscomb CE (2001) Mergers in the publishing industry. *Bull Med Libr Assoc* 89:307–308
- Lotka AJ (1926) The frequency distribution of scientific productivity. *J Wash Acad Sci* 16:317–324
- Malakoff D (2000) Librarians seek to block merger of scientific publishing giants. *Science* 290:910–911
- McNulty M (1979) *Publishing*, book. In: *ALA yearbook: a review of library events*, Vol 4. American Library Association, Chicago, IL, p 225–228
- Miller M (2004) Faculty Senate discusses journal fees. Available online at: <http://daily.stanford.edu/article/2004/2/6/facSenDiscussesJournalFees>
- Moed HF, Van Leeuwen TN (2005) *Citation analysis in research evaluation*. Springer, New York
- Morais R (2002) Double Dutch no longer. *Forbes Magazine* no. 10, 11th Nov
- Moses H III, Dorsey ER, Matheson DHM, Thier SO (2005) Financial anatomy of biomedical research. *JAMA* 294:1333–1342

- Nature (2005) Editorial: Not-so-deep impact. *Nature* 435: 1003–1004, 23 Jun
- Newman MEJ (2005) Power laws, Pareto distributions and Zipf's law. *Contemp Phys* 46:323–351.
- Payne D (2004) Canadian research compromised? *The Scientist Daily News*, Nov 24. Available online at: www.the-scientist.com/news/20041124/02
- Perakakis P, Taylor M, Buela-Casal G, Checa P (2006) A neuro-fuzzy system to calculate the journal internationality index. *Proc CEDI2005 Symp Fuzzy Logic and Soft Comp*, p 157–163. Available online at: http://damir.iem.csic.es/~michael/OA/perakakis_et_al_2005.pdf
- Pirsig RM (1974) *Zen and the art of motorcycle maintenance*. William Morrow, New York
- Redner S (1998) How popular is your paper? An empirical study of the citation distribution. *Eur Phys J B* 4(2): 131–134. Available online at: <http://arXiv.org/cond-mat/9804163>
- Reed Elsevier (2007) Press release: Reed Elsevier to exit the defence exhibitions sector. Available online at: www.reed-elsevier.com/index.cfm?articleid=2084
- Rehmann U (2004). Math journal price survey, based on AMS 2007 data. Available online at: www.mathematik.uni-bielefeld.de/~rehmann/BIB/AMS/Price_per_Page.html
- Russo M (2006) A world of opportunity and influence. *J Assoc Lab Autom* 11:A6–A7
- Schmid JA (2003) Cover the eyes of Lady Justice — an appeal for double-blind peer reviewing. *EMBO Rep* 4:734–736
- SPARC (Scholarly Publishing and Academic Resources Coalition) (2000) *Declaring independence: a guide to creating community-controlled science journals*. SPARC Publications, Washington
- Scully C, Lodge H (2005) Impact factors and their significance; overrated or misused? *Br Dent J* 198:391–393
- Seglen PO (1992) The skewness of science. *J Am Soc Inf Sci* 43:628–638
- Seglen PO (1997) Why the impact factor of journals should not be used for evaluating research. *BMJ* 314:497–502. Also available online at: <http://www.bmj.com/cgi/content/full/314/7079/497>
- Sher IH, Garfield E (1966) New tools for improving and evaluating the effectiveness of research. In: Yovits MC, Gilford DM, Wilcox RH, Staveley E, Lemer HD (eds) *Research program effectiveness, proceedings of the conference sponsored by the Office of Naval Research*, Washington, DC, Jul 27–29, 1965. Gordon and Breach, New York, p 135–146
- Smith JWT (2004) The importance of access to academic publications for the developing world and the implications of the latest developments in academic publishing, 15–17 September 2004. Presented at the ICCC Int Conf Computer Comm: Core Platform for the Implementation of the Computer Society. Available online at: <http://library.kent.ac.uk/library/papers/jwts/develop.htm>
- Stafford T, Atiyah M, Banerjee B and others (2007) Ethics before arms. *Times Higher Education*, Mar 2:15
- Stergiou KI, Tsikliras AC (2006) Underrepresentation of regional ecological research output by bibliometric indices. *Ethics Sci Environ Polit* 2006:15–17
- Suber P (2006). Lists related to the open access movement. Nov 12. Available online at: www.earlham.edu/~peters/fos/lists.htm#actions
- Sutherland J (1999) *Who owns John Sutherland?* London Review of Books 7, London
- Taubes G (1993) Measure for measure in science. *Science* 260:884–886
- Thomson ISI (2004) Editorial, The Thomson Scientific journal selection process. Available online at: <http://scientific.thomsonreuters.com/free/essays/selectionofmaterial/journalselection/>
- Tuck A (2003) Impact factors: a tool of the sterile audit culture: Stalked by accountants, unaware of pre-Internet work, how can young scientists thrive? *Nature* 424:14
- US Census Bureau (1999) *Statistical abstract of the United States, science and technology*. US Census Bureau, Washington, DC
- Vaidya JS (2005) V-index: a fairer index to quantify an individual's research output capacity. *BMJ* 331:1339-c–1340-c
- Van de Sompel H, Lagoze C (2000) The Santa Fe convention of the open archives initiative. *D-Lib Magazine*, 6(2). Available online at: www.dlib.org/dlib/february00/vandesompel-oai/02vandesompel-oai.html
- Van de Sompel H, Lagoze C (2002) Notes from the interoperability front: a progress report on the open archives initiative. In: *ECDL 2002 Proc 6th Eur Conf Res Adv Tech Digital Libraries*. Springer-Verlag, London, p 144–157
- Vinkler P (1986) Evaluation of some methods for the relative assessment of scientific publications. *Scientometrics* 10: 157–177
- Wager E, Parkin EC, Tamber PS (2006) Are reviewers suggested by authors as good as those chosen by editors? Results of a rater-blinded, retrospective study. *BMC Med* 4:13
- Waltz E (1998) *Information warfare principles and operations*. Artech House Publishers, London
- Weiss R (2007) Publishing group hires 'Pit Bull of PR'. *Washington Post*, Jan 26:19
- Weisstein EW (2001) What happened to *Mathworld*. Available online at: http://mathworld.wolfram.com/about/erics_commentary.html
- Wysocki B (2005) Peer pressure: scholarly journals premier status is diluted by web. *Wall Street Journal*, 23 May
- Zipf GK (1949) *Human behavior and the principle of least effort*. Addison-Wesley, Cambridge, MA

Editorial responsibility: Howard Browman, Storebø, Norway (until Nov. 5, 2007), and Konstantinos Stergiou, Thessaloniki, Greece

*Submitted: July 23, 2007; Accepted: January 21, 2008
Proofs received from author(s): May 7, 2008*