e-access

These letters form part of *Nature*'s current debate on access to the scientific literature.
For more examples, see ▸ **http://www.nature.com/nature/debates/e-access/index.html**.

Free online availability substantially increases a paper's impact

Sir — The volume of scientific literature far exceeds the ability of scientists to identify and use all relevant information. The ability to locate relevant research quickly will dramatically improve communication and scientific progress. Although availability varies greatly by discipline, more than a million research articles are now freely available on the web.

Here we investigate the impact of free online availability by analysing citation rates. Online availability of an article may not greatly improve access and impact without efficient and comprehensive search services; a substantial percentage of the literature needs to be indexed by these search services before scientists consider them useful. In computer science, a substantial percentage of the literature is online and available through search engines such as Google (www.google.com), or specialized services such as ResearchIndex (www.researchindex.org) — although the greatest impact of online availability is yet to come, because comprehensive search services and more powerful search methods have become available only recently.

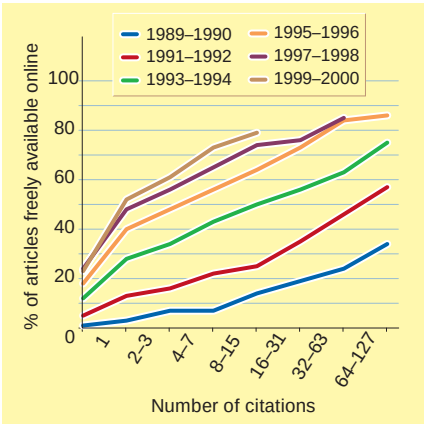
We analysed 119,924 conference articles in computer science and related disciplines, obtained from DBLP (dblp.uni-trier.de). In these fields, conference articles are typically formal publications and are often more prestigious than journal articles, with acceptance rates at some conferences below 10%. We estimated citation counts and online availability using ResearchIndex, excluding self-citations.

The figure shows the probability that an article is freely available online as a function of the number of citations to the article, and the year of publication of the article. The results are dramatic, showing a clear correlation between the number of times an article is cited and the probability that the article is online. More highly cited articles, and more recent articles, are significantly more likely to be online, in computer science. The mean number of citations to offline articles is 2.74, and the mean number of citations to online articles is 7.03, an increase of 157%.

We analysed differences within publication venues (the proceedings of a conference for a particular year, for example), looking at the percentage increase in citation rates for online articles. When offline articles were more highly cited, we used the negative of the percentage increase for offline articles: hence if the average number of citations for offline articles is two, and the average for online articles is four, the percentage increase would be 100%. For the opposite situation, the percentage increase would be −100%. Averaging the percentage increase across 1,494 venues containing at least five offline and five online articles results in an average of 336% (median 158%) more citations to online computer-science articles compared with offline articles published in the same venue.

If we assume that articles published in the same venue are of similar quality, then the analysis by venue suggests that online articles are more highly cited because of their easier availability. This assumption is likely to be more valid for top-tier conferences with very high acceptance standards. Restricting our analysis to the top 20 publication venues by average citation rate gives an increase of 286% (median 284%) in the citation rate for online articles.

Free online availability facilitates access in many ways, including provision of online archives; direct connections among scientists or research groups; hassle-free links from e-mail, discussion groups and other services; indexing by web search engines; and the creation of third-party search services. Free online availability of scientific literature offers substantial benefits to science and society. To maximize impact, minimize redundancy and speed scientific progress, authors and publishers should aim to make research easy to access.



Analysis of 119,924 conference articles in computer science and related disciplines. The actual percentage of articles available online is greater, owing to limitations in the extraction of article information from online documents and limitations in locating articles on the web. Only points with greater than 100 articles are computed.

Steve Lawrence
NEC Research Institute, 4 Independence Way, Princeton, New Jersey 08540, USA

Authors willing to pay for instant web access

Sir — Authors of journal articles want their efforts certified by peer review and made conveniently available to the widest possible readership. They do not expect or receive royalties. What they hope for is 'impact' — attention, especially from other researchers, and recognition, especially from those who decide on hiring and promotions. Journal articles have greater impact if they are immediately and widely accessible. Maximum impact is achieved by immediate free web access (see *Nature* **410,** 1024–1025; 2001).

Journal publishers are taking advantage of the web's quick, convenient delivery of information by creating electronic versions of their traditional print journals, accessible only to subscribers or to clients of an institution with a site licence. Authors are not offered immediate free web access for their articles.

The Entomological Society of America (ESA), publisher of four leading entomological journals, recently began selling immediate free access (see http://csssrvr.entnem.ufl.edu/~walker/tjwbib/walker.htm). The results suggest that a market-driven transition to free access for all articles in all journals is possible.

ESA's business plan is simple: it will provide immediate free web access, at a fair price, to authors who want it. As the cost of offering this rises (because of subscription cancellations), the price will increase. No author will be required to purchase it, and sales of subscriptions to the journals will continue as long as they are profitable. The endpoint of this plan is uncertain, but it may lead to the demise of paper publication and subscriptions, as authors and the institutions that support them embrace free access and strive to reduce costs.

Direct costs of the present system include printing paper issues, limiting access to electronic versions, and making past and present volumes accessible in hundreds of research libraries. Indirect costs are reduced impact of articles and severely restricted access by researchers in smaller institutions and in developing countries. Nonetheless, many stakeholders

believe that printed issues, or at least tolls in the form of subscriptions and site licences, will continue indefinitely.

ESA began selling immediate free web access in January 2000. During the first two months of the service, authors bought it for 13% of articles, rising steadily to 59% during March and April 2001. The price for the service is currently 75% of the price of 100 paper reprints, for example $90 for a 7-page article. This price provides a greater profit margin than for paper reprints, which are expensive to produce and deliver. Immediate free web access requires only that the PDF file of the article is made freely accessible on ESA's web server.

If immediate free access is a profit-making service that many authors want and will pay for, why is ESA apparently the only publisher that sells it? For scientific societies, the answer is probably that their institutional inertia is great and their members have yet to lobby for it — as ESA members did. Commercial publishers may fear that selling immediate free access to those who want it may lead to all authors buying it, in which case revenues from subscriptions and site licences might cease.

On the other hand, societies have supplemented modest incomes from lower-priced library subscriptions with member dues and page charges. Without journal subscriptions, societies and commercial publishers will collect page charges to pay for refereeing, editing and composing. Publishers will pay nothing to make the articles freely web accessible, because research libraries and PubMed Central will post them without charge.

Authors should encourage publishers to provide immediate free access at a fair price. Other things being equal, many will prefer to publish in journals that provide it, especially as electronic literature indexes begin linking directly to the e-versions of articles. Most authors would like nothing better than for their articles to be available in full text, without tolls, via links in widely used literature indexes.

Thomas J. Walker
Department of Entomology and Nematology, University of Florida, PO Box 110620, Gainesville, Florida 32611-0620, USA

Impact factors, and why they won't go away

Sir — Brunstein in Correspondence[1] asks whether online publishing will herald the end of impact factors. Unless he is forecasting the end of print publications altogether, this is doubtful. Were print journals to disappear, however, I am confident that a new impact factor would be invented. Information scientists are already computing web impact factors[2].

It would be more relevant to use the actual impact (citation frequency) of individual papers in evaluating the work of individual scientists rather than using the journal impact factor as a surrogate. The latter practice is fraught with difficulties, as Seglen and others have pointed out[3]. As long as scientists publish articles containing lists of cited references, it will be possible to calculate impact factors. It is to be hoped that citation practices on the web will become sufficiently standardized to permit accurate calculations.

It will be necessary to distinguish between citations to URLs for research articles, on the one hand, and, on the other, to readerships as reflected in 'webometric' studies measuring web activity. One ordinarily assumes that there are many more readers than citers, but there is a widespread mythology that authors are cited more than they are read!

Eugene Garfield
Institute for Scientific Information, 3501 Market Street, Philadelphia, Pennsylvania 19104, USA

1. Brunstein, J. *Nature* **403,** 478 (2000).
2. Bjorneborn, L. & Ingwersen, P. *Scientometrics* **50** (1), 65–82 (2001).
3. Seglen, P. O. *Br. Med. J.* **314,** 498–502 (1997).

Evolution is what's needed, not revolution

Sir — In the current debate on web availability of peer-reviewed scientific literature there seems little attempt to retain the advantages of the present system of journals.

These are, first, the element of specialization (the journal's field) and, second, the quality and novelty of the science the journal demands for publication (its place in the pecking order).

Both aspects are crucial, as authors wish their articles to be immediately read by the appropriate audience and to accrue the kudos appropriate to the significance of their work. So any universal web-based system must accommodate these two aspects.

This could be achieved if existing journals were replaced by equivalent websites that were run not by commercial publishers but by the learned societies on a non-profit basis. Provided the societies did not use their sites to generate extra income, they could keep costs to a minimum.

Reviewing could be carried out as now; the costs would be covered by page charges to authors (for both accepted and rejected submissions), and by small charges for web access to published articles. The subscription charges would be scaled: relatively high for an institution subscribing for the benefit of all its staff, at intermediate level for an individual laboratory, and very low for a personal subscription.

Control of access could be through the predesignated IP addresses of the servers and individual terminals of subscribers. Thus an individual laboratory could have online subscription to its favourite sites for a fraction of the cost that it must now pay to get them as printed journals. If a library wanted paper on its shelves, it could either pay the site to sell it what we currently call a 'journal' or it could have a licence to print the content out as part of its institutional subscription.

A great deal of the effort for web publication (for example, generating PDF versions of text and figures in the house style of the site) can be undertaken by the authors, since they are the keenest to see the article in the public domain.

The current requirements for publication in the *Journal of Biological Chemistry*, in which every aspect is electronic, show that this is a straightforward procedure, which, while not cost-free, is not prohibitively expensive when printing costs do not have to be covered.

Colyn Crane-Robinson
Biophysics Laboratories, University of Portsmouth, St Michael's Building, White Swan Road, Portsmouth PO1 2DT, UK

The Net is many people's only chance of access

Sir — Pakistan is not on the publishing map. I doubt if the collective population of 140 million manage to subscribe to more than a few costly journals. On the other hand we have only about 300,000 computer users. Withholding full text from a country such as Pakistan is thus ridiculous. I would suggest that full text should be made available to everyone on the Internet. This would not affect journals financially, since most people using this service in countries such as Pakistan would never be able to subscribe to them.

F. A. Khan
1A White House Lane, Sunder Dass Road, Lahore, Pakistan

Erratum In Stevan Harnad's Commentary on freeing the scientific literature (*Nature* 410, 1024–1025; 2001), the estimate of the minimum cost of peer review was not from the American Institute of Physics but from a summary of a group discussion by Mark Doyle of the American Physical Society. The $500 estimate used in that discussion included only peer-review costs, not post-acceptance costs. The URL for the estimate is: http://documents.cern.ch/archive/electronic/other/agenda/a01193/a01193s5t11/transparencies/.