Title
Emerging alternatives to the impact factor

Journal
OCLC Systems & Services: International Digital Library Perspectives, 24(3)

Authors
Banks, Marcus A.
Dellavalle, Robert

Publication Date
2008

Peer reviewed

Emerging Alternatives to the Impact Factor

General Review

Keywords: impact factor; h index; Y factor; evaluation metrics

Purpose:

The authors document the proliferating range of alternatives to the impact factor that have arisen within the past five years, coincident with the increased prominence of open access publishing.

Methodology/Approach:

This paper offers an overview of the history of the impact factor as a measure for scholarly merit; a summary of frequent criticisms of the impact factor's calculation and usage; and a framework for understanding some of the leading alternatives to the impact factor.

Findings:

This paper identifies five categories of alternatives to the impact factor:

a. Measures that build upon the same data that informs the impact factor.
b. Measures that refine impact factor data with "page rank" indices that weight electronic resources or Web sites through the number of resources that link to them.
c. Measures of article downloads and other usage factors.

d. Recommender systems, in which individual scholars rate the value of articles and a group's evaluations pool together collectively.
e. Ambitious measures that attempt to encompass the interactions and influence of all inputs in the scholarly communications system.

Value of Paper:

Librarians can utilize the measures described in this paper to support more robust collection development than is possible through reliance on the impact factor alone.

History and Calculation of the Impact Factor

In 1955 Eugene Garfield proposed a "bibliographic system for science literature that can eliminate the uncritical citation of fraudulent, incomplete, or obsolete data by making it possible for the conscientious scholar to be aware of criticisms of earlier papers" (Garfield 1955). Using the example of legal documents, which heavily depend on the precedent of previous rulings, Garfield proposed development of an "impact factor" that would rate the worth of a scientific article based upon the number of articles that cite it subsequently. The impact factor would also be useful for tracing the "eclectic" connections between ideas that would not be obvious through more established methods like subject indexing. Finally, provided with a convenient way to access a paper's subsequent citations, the "conscientious scholar" would be in a better position to build upon the chain of reaction to any given paper.

In 1961 Garfield and Sher began to operationalize these goals, with the creation of the *Science Citation Index* (SCI) (Garfield 2006). Although the SCI focused upon the impact factor for authors, by the end of the decade Garfield became interested in quantifying the impact factor for journals (Garfield 1972). The journal-focused metric eventually became quantified in the *Journal Citation Reports* (JCR) database. Anyone who inquires about the "high impact journals in my field" probably wants data from the JCR.

The impact factor for a journal "x" is a ratio based upon the previous two years of citation data [1]. As an example, the 2007 impact factor for "x" is calculated as follows:

$$\frac{\text{Number of 2007 articles in journals that reference citable articles published in "x" in 2005-2006}}{\text{Number of all citable articles published in "x" in 2005-2006}}$$

If the numerator is 200, and the denominator is 100, then the impact factor for "x" would be 2.

In general, the higher impact factor journals in a given field are perceived to be more prestigious than other titles.

Crucially, the impact factor's calculation depends on the number of "citable items" a journal published. Original research papers are always included, but letters and editorials often are not. This stipulation alone can have a significant effect on a given journal's impact factor.

As of 2007, the journal impact factor remains a foremost metric in scholar's minds. Everyone wants their work to appear in high impact journals. Librarians must respond to this impulse in collection decisions, by privileging these journals even if they are very expensive to obtain.

Criticism of the Impact Factor

Well before the Internet created space for open access publishing, criticisms of the impact factor abounded. Seglen summarizes many of these critiques in a 1997 paper that built upon earlier work by himself and others (Seglen 1997). Among other problems, Seglen concludes that a relatively small number of highly cited articles can disproportionately skew the impact factor for a journal; that review articles are cited so frequently that this favorably influences the impact factors for journals that publish many review articles; and that the impact factor arbitrarily favors research in fields whose literature rapidly becomes obsolete.

More recent criticism by Smith emphasizes the deleterious role of "citation cartels" in which authors indiscriminately cite themselves and others in order to boost the impact factor (Smith 2006). Another negative consequence of obeisance to the impact factor is the devaluing of "softer" elements of a journal—such as editorials and layman's overviews—that are less likely to be cited by researchers. Citing Malcolm Chiswick, Smith proclaims that the result of this diminution is an "impacted journal." The editors of *PLoS Medicine* brand all such maneuvers as playing the "impact factor game" (PLoS Medicine Editors 2006).

In a recent lecture, Garfield defends the impact factor while acknowledging some shortcomings in its calculation (Garfield 2006). He argues that journal impact factors for discrete fields tend to be stable over time, which validates their usefulness as an indicator of the most prestigious journals. He also charges that critics of the impact factor exaggerate the effects of anomalies in its calculation. With those lines drawn, Garfield admits that the exclusion of non-citable items like editorials can mildly distort the impact factor (although he only admits this for high-profile medical titles.) Next Garfield describes his continuing work to improve the value of the impact factor. He then points readers to another database, *Journal Performance Indicators*, which "eliminate[s]" "many of the discrepancies inherent" in traditional impact factor calculation. But although he acknowledges some defects, naturally Garfield is partial to his creation. He concludes with a quote from Hoeffel: "'Experience has shown that in each specialty the best journals are those in which it is most difficult to have an article accepted, and these are the journals that have a high impact factor.'"

The criticism of the impact factor has gained some traction. In 1997, Seglen was particularly concerned about the prospect that national governments would unfairly utilize impact factor scores as a convenient metric for evaluating scholars; those who published in high impact journals would reap future grants, while those who did not would go without. Smith echoed this fear in 2006. Both authors should be pleased with this statement from the British government about the UK's 2008 Research Assessment Exercise: "No panel will use journal impact factors as a proxy measure for assessing quality" (Research Assessment Exercise Team 2006).

Alternatives to the Impact Factor

A. Measures that build upon the same data that informs the impact factor

Proposed by Hirsch in 2005, the *h* index is computed as, "the number of papers with citation number ≥*h*" (Hirsch 2005). An author with an *h* index of 30 has published 30 papers that received at least 30 citations in subsequent work. If one scholar has published 100 papers and another has published 35, but both have 30 papers that received at 30 citations, each would have an *h* index of 30. Hirsch argues that his index provides a way to gauge the relative accomplishments of different researchers that is more rigorous than simply toting up their number of publications.

Hirsch describes calculating the *h* index using the data in the "times cited" field for author records in Thomson ISI's Web of Science database; this is the same data source that informs the calculation of a journal's impact factor. Although the *h* index is relatively straightforward, Hirsch recognizes that it should only be one factor in any evaluation of a scholar's impact. Almost as soon as Hirsch proposed the *h* index, researchers began to develop variants of it. A recent review summarizes the analytic potential of nine *h* index variants in the field of biomedicine (Bornmann et al. 2008).

Barendse recently built upon the *h* index in his proposal for a "strike rate index" (Barendse 2007). Using logarithmic techniques to ensure that journals in different fields can be compared objectively, Barendse transforms the *h* index from a strictly author-focused measurement into one that can also inform the evaluation of journals. This is analogous to Garfield's shift in

emphasis from authors to journals almost 40 years ago. However, Barendse argues that the strike rate index controls for citation patterns over the long term much better than the native impact factor. He finds the two approaches to be complementary—not competitive—means of understanding scholarly worth.

B. Measures that refine the impact factor with "page rank" data

In 2006 Bollen and colleagues introduced the "Y factor," a measure that combines traditional impact factor data with "page rank" data similar to what Google uses in its search algorithm (Bollen et al. 2006). To some extent, Google simply keeps track of which Web sites link to other Web sites. This is akin to rudimentary article citation. However, Google also factors the prestige of a linking site into its calculations; if the National Institutes of Health site links to a Web page, this means more than a link to the same page from a personal blog. The Y factor—which the authors acknowledge has not yet been fully justified scientifically—applies the page rank approach to journal citation networks. They found significant differences between the highest impact factor journals and Y factor journals in physics, computer science, and medicine, and less difference in a subspecialty of medicine, dermatology (Dellavalle et al. 2007).

Developed by Bergstrom, the Eigenfactor provides an online suite of tools that "ranks journals much as Google ranks websites" [2]. Available at no charge, the Eigenfactor attempts to account for the prestige of citing journals; incorporates many non-standard items such as newspapers and PhD dissertations into the citation network; and evaluates items over a 5 year (rather than 2 year) period. Of particular interest to librarians, the "cost-effectiveness search" relates this data to the

going subscription rates for journals as means of determining value-for-money. The "Article Influence" metric within the Eigenfactor is comparable to the impact factor, but that is just one aspect of the broader framework.

C. Measures of article downloads and other usage factors

The alternatives to the impact factor described thus far are refinements and enhancements of it. However, the profusion of online journal content in recent years has prompted calls for new evaluation metrics entirely; the impact factor's origin dates from a period when all scholarship was in print. The Usage Factor, which is still in its formative stages, is one alternative currently receiving attention. Its calculation relies upon statistics generated by COUNTER, which tracks monthly full-text article requests for many scholarly journals [3]. Internet usage data is often obtained from the web site Alexa [4].

The Usage Factor would be calculated as follows (Shepard 2007):

$$\frac{\text{Total usage of a journal (Counter journal usage for a specified period)}}{\text{Total number of articles published online during the same period}}$$

An even more straightforward alternative metric for the impact of an article is the number of times it has been downloaded, which is easy to track online. For the fields of physics, astrophysics, and mathematics, Brody and colleagues found a positive correlation between the frequency of times an article is downloaded soon after publication and the eventual rate at which

it will be cited (Brody et al. 2006). The attention to downloads refocuses evaluation on individual papers and away from journals. This is a positive development in light of the pernicious incentives that high impact journals can provide for author behavior.

Because open access articles are available to anyone with an Internet connection, they are likely to be downloaded more frequently than articles that are only available via personal or library subscriptions. Hitchcock continues to update a useful bibliography about the relationship between downloads and citation impact (Hitchcock 2007). As a practical matter, the open access download-to-citation advantage may be small (Kurtz et al. 2007); this is a topic for continuing debate. But it does seem evident that tracking downloads will be important to the evaluation of scholarly impact for the foreseeable future. In a 2005 survey—which has somewhat limited utility, due to the small sample size—established researchers rated downloads slightly higher than citations as an indicator of the impact of an article (Rowlands et al. 2005).

D. Recommender systems and other communitarian metrics

It is possible to track article downloads and calculate usage factors for all articles, open access or not. However, several community-based measurements are emerging that are much more practical in an open access environment.

At a 2003 symposium sponsored by the National Academies, Resnick promoted the concept of "recommender systems" as a means of expanding the peer review process (Resnick 2004). Rather than cloaking the entire process in secrecy, Resnick argued that journals should post the

contents of peer reviews, and possibly even the reviewer's names as a means of increasing transparency and ensuring that reviews were more thorough. In this more public context, one's colleagues would be encouraged evaluate the cumulative value of your peer reviews. Your prestige would rise (or fall) accordingly.

PLoS ONE, one of the most innovative open access journals published by the Public Library of Science, follows the spirit if not the letter of Resnick's proposals [5]. Papers published in PLoS ONE must pass an initial threshold review by an editor, but most of the reviewing (and rating) occurs publicly after an article appears. The home page for the journal provides easy access to recent ratings—on the dimensions of insight, reliability, and style—and reader comments. The entire enterprise depends on "openness" in a full sense, from the author's willingness to receive public scrutiny to a reviewer's willingness to sign their name.

Schnell recently described an even more novel evaluation metric, which he termed the "Blog Citation Index" or BCI (Schnell 2007). An increasing (but still small) amount of scholarly discourse occurs via blog posts; the comments on these posts; and the links to the posts from other sites (which are akin to citations). The BCI would formally track these connections. The blogosphere is a radically open place that has blossomed in recent years; Schnell's proposal is a natural extension of earlier calls for "recommender systems" to enrich peer review.

E. Package of evaluation metrics designed for modern scholarly communication practices

Funded by the Andrew W. Mellon Foundation, the MESUR (MEtrics from Scholarly Usage of Resources) project is a two year effort to enrich "the toolkit used for the assessment of the impact of scholarly communication items, and hence of scholars, with metrics that derive from usage data" [6]. Researchers at the Los Alamos National Laboratory—who also developed the Y factor—are leading the MESUR (pronounced "measure") initiative. MESUR is the most comprehensive effort to date to align article impact evaluation techniques with modern scholarly communication practices, which are very different today than just a decade ago. The final report is due to Mellon in October 2008.

Conclusion

The impact factor is a scholarly evaluation metric that continues to have some utility. But it was conceived in a very different context than exists today, and was controversial even before the birth of the Internet and the advent of open access publishing fundamentally altered the landscape. The impact factor is a tool whose usefulness is waning, but there is not yet a fully viable alternative to it. This paper describes several alternatives to the impact factor, some well-developed and some still in formative stages. Librarians should inform themselves about the strengths and weaknesses of the alternatives; contribute to the discussion about their ongoing development; encourage and facilitate faculty awareness of these alternatives; and begin to use them in deciding how to develop and promote scholarly resources.

References

1. Barendse, W. (2007) “The strike rate index: a new index for journal quality based on journal size and the h-index of citations”, Biomedical Digital Libraries, Vol 4 No 3. Available http://www.bio-diglib.com/content/4/1/3.

2. Bollen, J., Rodriguez, M.A., Van de Sompel, H. (2006) “Journal status,” Scientometrics, Vol 69 No 3, pp. 669-687. Available http://arxiv.org/abs/cs/0601030.

3. Bornmann, L., Mutz, R., Daniel, H-D. (2008) “Are there better indices for evaluation purposes than the *h* index? A comparison of nine different variants of the *h* index using data from biomedicine,” Journal of the American Society for Information Science and Technology, Vol 59 No 5, pp. 830-837.

4. Brody, T., Harnad S., Carr L. (2006) “Earlier Web usage statistics as predictors of later citation impact,” Journal of the American Association for Information Science and Technology, Vol 57 No 8, pp 1060-1072. Available http://eprints.ecs.soton.ac.uk/10713/.

5. Dellavalle, R.P., Schilling, L.M., Rodriguez, M.A., Van de Sompel, H., Bollen, J. (2007) “Refining dermatology impact factors using PageRank,” Journal of the American Academy of Dermatology Vol 57 No 1, pp. 116-119.

6. Garfield, E. (2006) “The history and meaning of the journal impact factor,” Journal of the American Medical Association Vol 295 No 1, pp. 90-93. Available http://jama.ama-assn.org/cgi/content/full/295/1/90.

7. Garfield, E. (1972) “Citation analysis as a tool in journal evaluation,” Science Vol 178 No 60, pp. 471-479.

8. Garfield, E. (1955) “Citation indexes for science: a new dimension in documentation through association of ideas,” Science Vol 122 No 3159, pp. 108-11. Available http://ije.oxfordjournals.org/cgi/content/full/35/5/1123.

9. Hirsch, J. E. (2005) “An index to quantify an individual’s scientific research output,” Proceedings of the National Academy of Sciences Vol 102 No 46, pp. 16569-16572. Available http://www.pnas.org/cgi/content/full/102/46/16569.

10. Hitchcock, S. (2007) “Effect of open access on citation impact: a bibliography of studies.” Available http://opcit.eprints.org/oacitation-biblio.html.

11. Kurtz, M.J., Henneken, E.A. (2007) “Open access does not increase citations for articles from The Astrophysical Journal.” Available http://arxiv.org/abs/0709.0896.

12. Public Library of Science Medicine Editors (2006) “The impact factor game,” Public Library of Science Medicine Vol 3 No 6, pp. e291. Available http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=16749869.

13. Research Assessment Exercise Team (2006) “RAE2008: panel criteria and working methods.” Available http://www.rae.ac.uk/pubs/2006/01/.

14. Resnick, P. (2004) "Implications of Emerging Recommender and Reputation Systems," in Committee on Electronic, Scientific, Technical, and Medical Journal Publishing (ed.), Electronic Scientific, Technical and Medical Journal Publishing and its Implications: Report of a Symposium, Washington, pp.49-50.

15. Rowlands, I., Nichols, D. (2005) "New journal publishing models: an international survey of senior researchers." Available http://www.ucl.ac.uk/ciber/ciber_2005_survey_final.pdf.

16. Schnell, E. (2007) "A blog citation index?" Available http://ericschnell.blogspot.com/2007/11/blog-citation-index-bci.html.

17. Seglen, P.O. (1997) "Why the impact factor of journals should not be used for evaluating research," British Medical Journal Vol 314 No 7079, pp. 498-502. Available http://www.bmj.com/cgi/content/full/314/7079/497.

18. Shepard, P.T. (2007) "Final report on the investigation into the feasibility of developing and implementing journal usage factors." Available http://www.uksg.org/sites/uksg.org/files/Final%20Report%20on%20Usage%20Factor%20project.pdf.

19. Smith, R. (2006) "Commentary: The power of the unrelenting impact factor—is it a force for good or harm?," International Journal of Epidemiology Vol 35 No 5, pp. 1129-1130. Available http://ije.oxfordjournals.org/cgi/content/full/35/5/1129.

Web Sites

1. “The Thomson Scientific Impact Factor.” Available http://scientific.thomson.com/free/essays/journalcitationreports/impactfactor/.

2. “Why Eigenfactor.org: Ranking and Mapping Scientific Journals.” Available http://www.eigenfactor.org/whyeigenfactor.htm.

3. “COUNTER: Online Usage of Electronic Resources.” Available http://www.projectcounter.org/.

4. “Alexa the Web Information Company.” Available http://alexa.com/.

5. “PLoS ONE: Publishing science, accelerating research.” Available http://www.plosone.org/home.action.

6. “Mesur: Metrics from Scholarly Usage of Resources.” Available http://mesur.org/Home.html.