Peer Review, Citation Ratings and Other Fetishes

Jennifer A. Foley

Received: 8 November 2012 / Revised: 9 January 2013 / Accepted: 22 January 2013 / Published online: 6 February 2013
© Springer International Publishing AG 2013

Abstract Academic success or ability is often assessed through peer review or the seemingly more objective method of citation ratings. However, citation ratings may be more objective in that they offer a more automated or mechanical method of assessing quality, but it does not necessarily follow that these ratings provide an assessment of genuine scientific impact. Likewise, although peer review may provide an effective filtering system, it cannot be assumed that it provides objective critiques. This review discusses the fetishes and flaws of both methods, and suggests that future reviewing methods should involve both quantitative and qualitative methods, tailored to the specific individual subject area.

Key words Citations · Peer review · Impact factor · *h*-index

It is argued that academic success or quality should be assessed using a combination of both seemingly objective indicators, such as citation counts or single number indices like the *h*-index, and more subjective, or expertise-informed, peer reviews [38]. These measures are often used to inform decisions about grant/institutional funding, staff hiring, promotion and/or tenure [1, 26]. However, both of these ostensible indicators of quality are criticized for having significant weaknesses, and appear open to significant bias and manipulation.

Since the 17th century, when the first ‘scientific’ journals began, manuscripts submitted to journals for publication have usually been subject to peer review. In this process, the journal editor usually consults with other experts, or ‘peers’, to provide an objective and skilled critique of the science presented [27]. This process allows editors to screen for any flaws, and select the best papers for publication [22, 43]. In some cases, peer reviewers may also act as ‘shepherds’, working with the authors to improve the quality of a paper [2, 8]. Therefore, peer review is often discussed as a sacred process, the linchpin of scientific rigour, ensuring the quality of published papers [43].

However, peer reviews are subject to a number of different influences and rarely completely objective [36]. Indeed, reviewers have been found to give more favourable reviews to experts in the field [41, 44], who are from more prestigious institutions [28, 29, 37, 41], are male [44] and speak English as their first language [39].

Many of these issues may be obviated using blinded reviews (‘single’ blinded when the authors are anonymous, and ‘double’ blinded when both authors and reviewers are anonymous [29, 32]). Indeed, one study reported a significant increase in the number of female first-authored papers published in one journal following the adoption of double-blinded reviews [12]. However, double-blinded reviews have also been criticized for preventing the reviewer from judging the novelty of the contribution (does the paper represent true scientific advance, or the rehashing of the author’s previously published data?) or the author’s level of expertise, leading to possible inappropriate reviewer comments [33].

Furthermore, double-blinded reviews may not be truly anonymous. An expert reviewer may actually be able to identify the author through self-citations and references to earlier works. Similarly, authors may be able to identify the identity of the reviewer through the nature of their

J. A. Foley (✉)
Department of Neuropsychology, National Hospital for
Neurology and Neurosurgery, Queen Square,
London, WC1N 3BG, UK
e-mail: Jennifer.Foley@uclh.nhs.uk

comments, thus reducing the effectiveness of any such blinding procedures.

There are also issues around conflict of interests, when reviewers may wish to delay or suppress certain work being published, or indeed lack specialist knowledge, failing to recognize either errors or important contributions. Thus, editors may seek the advice of a number of reviewers, perhaps chosen for their differing areas of expertise. However, each review can take considerable time, and a greater number of authors may lead to a greater likelihood of disagreement between reviewers, or lack of consistency, with little evidence that an increased number of reviewers leads to better papers [9].

Despite these shortcomings, several studies have found peer reviews to have high predictive validity, with a high association between reviewers' ratings and later citation ratings of both individual papers [10, 35] and individual scientists [9].

Citation ratings are the number of times that a paper has been cited in published work, and thus a higher number of citations is thought to reflect more important or influential work. Citation ratings are also used to calculate the personal impact of individual scientists, through single number indices, such as the *h*-index [24]. The *h*-index is calculated by determining the number of papers a scientist has with at least *h* citations each, with other papers having no more than *h* citations each. Such single number indices have been applauded for allowing the quick calculation of an individual's impact [7], with clear implications for that individual.

However, citation ratings can be misleading. Citation ratings assume that authors reference all the previous work that has influenced them, but of course authors often have limited space in which to include citations, and thus must choose to include only a select few, and face competing pressures when doing so [41]. Several studies have found that authors tend to cite papers that have a greater number of authors [18, 42], are longer in length [41], are in a journal issue that contains a high-impact article [25], are reviews [3] and report positive findings [31]. Moreover, citations may be negative, where papers are cited to criticize work, rather than praise [21].

Citation ratings may also be manipulated. Authors may make reference to their own articles, inadvertently artificially inflating their apparent impact [14, 19], which may be particularly significant if the paper has multiple authors [5]. Schreiber [40] argues that self-references should be removed from citation ratings; but, self-references have also been found to increase the number of citations from others [20] and, therefore, solely subtracting the number of self-references from citation rates may not actually remove their total effect, but also handicap productive groups [19].

Citation rates are also used to calculate journal impact factors, with higher impact factors often taken as a proxy of journal quality. Indeed, journals with higher citation rates

have lower acceptance rates, suggesting that such journals publish better quality papers [13, 23]. The most widely used method for calculating impact factor is that published by Thomson Reuters. In this calculation, the total number of citations gathered during a specific year by articles published in the previous 2 years is divided by the number of substantive or 'source' items. Total number of citations encompasses citations to all items published within the journal, but the number of source items includes only articles, reviews and proceeding papers. Thus, it is clearly possible to massage impact factor by reducing the number of source items, and boosting the number of other types of publications, such as letters, serving to increase the overall citation rating. Journal articles also tend to have an unequal distribution of citations, with some articles receiving more attention than others [11]. Such intra-journal variation is obscured through the current calculation process.

Citation rates also differ according to subject area, with certain disciplines being faster moving than others [15–17], and certain topics more attractive for publication in wider science journals. For instance, fundamental science journals appear to have larger mean impact factors than do journals of specialized, or applied, subject areas [5], and thus it is not possible to compare citations or impact factor across subject areas. Citation rates also vary according to whether or not the journal is open-access, with freely accessible papers gathering more citations [6, 30, 34].

In sum, citation ratings are often thought to be more of an 'objective' indicator of quality than 'subjective' peer review, but is this only an illusion? Citation ratings may be more objective in that they offer a more automated or mechanical method of assessing quality, but it does not necessarily follow that these ratings provide an assessment of genuine scientific impact. Likewise, although peer review may provide an effective filtering system, it cannot be assumed that it provides objective critiques. Indeed, it appears that although both methods are lauded, both methods are flawed.

Therefore, journals are increasingly adopting novel techniques to assess quality. These include conducting public as well as traditional peer reviews; asking reviewers to rank rather than review papers; asking for papers to be submitted with any previous reviewer comments; as well as ranking reviewers themselves [4, 8]. The jury is still out on whether any of these newer methods actually improve the assessment of scientific quality, but perhaps a combination of both quantitative and qualitative methods, tailored to each subject area, would offer the best compromise.

References

1. Adler KB (2009) Impact factor and its role in academic promotion. *Am J Respir Cell Mol Biol* 41:127

2. Adler R, Ewing J, Taylor P (2009) Citation statistics: a report from the international mathematical union (IMU) in cooperation with the international council of industrial and applied mathematics (ICIAM) and the institute of mathematical statistics (IMS). *Stat Sci* 24:1–14
3. Aksnes DW (2006) A macro study of self-citation. *Scientometrics* 56:235–246
4. Akst J (2010) I hate your paper. *The Scientist* 24:36
5. Amin M, Mabe M (2000) Impact factors: use and abuse. *Perspect Publ* 1:1–6
6. Antelman K (2004) Do open-access articles have greater impact? *Coll Res Libr* 65:372–382
7. Ball P (2005) Index aims for fair ranking of scientists. *Nature* 436:900
8. Birukou A, Wakeling JR, Bartolini C, Casati F, Marchese M, Mirylenka K, Osman N, Ragone A, Sierra C, Wassef A (2011) Alternatives to peer review: novel approaches for research evaluation. *Front Comput Neurosci* 5:1–12
9. Bornmann L, Daniel H-D (2005) Selection of research fellowship recipients by committee peer review: reliability, fairness and predictive validity of board of trustees' decisions. *Scientometrics* 63:297–320
10. Bornmann L, Daniel H-D (2010) The validity of staff editors' initial evaluations of manuscripts: a case study of *Angewandte Chemie international edition*. *Scientometrics* 85:681–687
11. Buchtel HA, Della Sala S (2006) Impact factor: does the 81/20 rule apply to Cortex? *Cortex* 42:1064–1065
12. Budden AE, Tregenza T, Aarssen LW, Koricheva J, Leimu R, Lortie CJ (2008) Double-blind review favours increased representation of female authors. *Trends Ecol Evol* 23:4–6
13. Buffardi LC, Nichols JA (1981) Citation impact, acceptance rate, and APA journals. *Am Psychol* 36:1453–1456
14. Della Sala S, Brooks J (2008) Multi-authors' self-citation: a further impact factor bias? *Cortex* 44:1139–1145
15. Della Sala S, Crawford JR (2006) Journal impact factor as we know it handicaps neuropsychology and neuropsychologists. *Cortex* 42:1–2
16. Della Sala S, Crawford JR (2007) A double dissociation between impact factor and cited half-life. *Cortex* 43:174–175
17. Della Sala S, Grafman J (2009) Five-year impact factor. *Cortex* 45:911
18. Figg WD, Dunn L, Liewehr DJ, Steinberg SM, Thurman PW, Barrett JC, Birkinshaw J (2006) Scientific collaboration results in higher citation rates of published articles. *Pharmacotherapy* 26:759–767
19. Foley JA, Della Sala S (2010) The impact of self-citation. *Cortex* 46:802–810
20. Fowler JH, Aksnes DW (2007) Does self-citation pay? *Scientometrics* 72:427–437
21. Garfield E (1979) Is citation analysis a legitimate evaluation tool? *Scientometrics* 1:359–375
22. Godlee F, Gale CR, Martyn CN (1998) Effect on the quality of peer review of blinding reviewers and asking them to sign their reports: a randomized control trial. *J Am Med Assoc* 280:237–240
23. Haensley PJ, Hodges PE, Davenport SA (2008) Acceptance rates and journal quality: an analysis of journals in economics and finance. *J Bus Finance Libr* 14:2–31
24. Hirsch JE (2005) An index to quantify an individual's scientific research output. *Proc Natl Acad Sci USA* 102:16569–16573
25. Hudson J (2007) Be known by the company you keep: citations—quality or chance? *Scientometrics* 71:231–238
26. Hyland K (2003) Self-citation and self-reference: credibility and promotion in academic publication. *J Am Soc Inform Sci Technol* 54:251–259
27. Ingelfinger FJ (1974) Peer review in biomedical publication. *Am J Med* 56:686–692
28. Judge TA, Cable DM, Colbert AE, Rynes SL (2007) What causes a management article to be cited—article, author, or journal? *Acad Manag J* 50:491–506
29. Laband DN, Piette MJ (1994) A citation analysis of the impact of blinded peer review. *J Am Med Assoc* 272:147–149
30. Lawrence S (2001) Free online availability substantially increases a paper's impact. *Nature* 411:521
31. Mahoney MJ (1977) Publication prejudices: an experimental study of confirmatory bias in the peer review system. *Cogn Ther Res* 1:161–175
32. McNutt RA, Evans AT, Fletcher RH, Fletcher SW (1990) The effects of blinding on the quality of peer review: a randomized trial. *J Am Med Assoc* 263:1371–1376
33. Nature (2008) Working double-blind. *Nature* 451:605–606
34. Norris M, Oppenheim C, Rowland F (2008) The citation advantage of open-access articles. *J Am Soc Inform Sci Technol* 59:1963–1972
35. Ophhof T, Coronel R, Janse MJ (2002) The significance of the peer review process against the background of bias: priority ratings of reviewers and editors and the prediction of citation, the role of geographical bias. *Cardiovasc Res* 56:339–346
36. Pendlebury DA (2009) The use and misuse of journal metrics and other citation indicators. *Scientometrics* 57:1–11
37. Peters DP, Ceci SJ (1982) Peer-reviewed practices of psychological journals: the fate of published articles, submitted again. *Behav Brain Sci* 5:187–195
38. Research Excellence Framework (2009). http://www.hefce.ac.uk/pubs/hefce/2009/09_38/#exec. Accessed 24 Aug 2012
39. Ross JS, Gross CP, Desai MM, Hong Y, Grant AO, Daniels SR, Hachinski VC, Gibbons RJ, Gardner TJ, Krumholz HM (2006) Effect of blinded peer review on abstract acceptance. *J Am Med Assoc* 295:1675–1680
40. Schreiber M (2007) Self-citations for the Hirsch index. *Europhys Lett* 78:1–6
41. Seglen PO (1998) Citation rates and journal impact factors are not suitable for evaluation of research. *Acta Orthop Scand* 69:224–229
42. Smart JC, Bayer AE (1986) Author collaboration and impact: a note on citation rates of single and multiple authored articles. *Scientometrics* 10:297–305
43. Smith R (2006) Peer review: a flawed process at the heart of science and journals. *J R Soc Med* 99:178–182
44. Wennerås C, Wold A (1997) Nepotism and sexism in peer-review. *Nature* 387:314–343