

## Citations and Ethics

Jan Reedijk\*



Jan Reedijk

Once upon a time, citations were used to inform the reader of a scientific paper which earlier research papers had been used for the described research and for the writing of the actual paper. All references were stored in a database that could be searched manually by browsing the Science Citation Index from the Institute of Scientific Information (ISI). These references have been used by scientists to efficiently study earlier scientific literature on a given topic and to see their follow-ups (the snowball effect). Nowadays, literature citations in a paper may also have a strong strategic content, like mentioning certain papers, for example, of potential referees, or ignoring others. Indeed, citation habits have changed, as deduced from complaints that some scientists, for example, disproportionately cite more, or less, papers from the same or the other side of the Atlantic or Pacific Ocean. Mentioning just one reference followed by the sentence “and references cited therein” is not uncommon. Referee reports requesting to add one or more “specific” references are also regularly seen.

Developments in recent years allow searching and consulting references to be performed by electronic analysis (usually online) that covers all literature published in scientific journals. ISI (now Thomson Reuters) has the most complete database, that is, the Web of Science (WoS, covers some 16 000 journals); Elsevier (Scopus) claims complete coverage after 1995 for 18 000 journals,

whereas Google Scholar as yet is the least complete database for this citation purpose. Availability of such databases and online searches may result in strategic and even unethical behavior.

### Rankings Based On Citation Analysis

It was Eugene Garfield, founder of ISI, who realized that the Science Citation Index database could be used for purposes other than literature retrieval, and he introduced the Journal Impact Factor (now called IF) in 1972<sup>[1]</sup> as a tool to monitor journals. Use of these (two-year) impact factors<sup>[\*]</sup> started to become fashionable in the 1990s, and the IF is till now THE parameter by which journals are ranked. The current IFs and the corresponding rankings are released each year in the Journal Citation Reports (JCR) from WoS.<sup>[\*\*][\*\*\*]</sup>

Recently Thomson Reuters performed and published analyses of the over 5000 most highly cited scientists in some 21

fields (<http://highlycited.com/browse>); searches for hot papers—for the last year, or the last decade—and for highly cited scientists can now also be performed by using other services of the Web of Science.

More recently, the Web of Science started to offer a fast online facility to determine the *h* indices for (groups of) persons—provided their complete and correct list of papers has been found and marked in the database.<sup>[\*]</sup> Initially, qualified bibliometric researchers analyzed the Thomson Reuters databases—often criticized by scientists in the early years—and made analyses and rankings of persons and institutions. Later, with the availability of ever more user-friendly interfaces, university and government administrators—not hindered by knowledge of citation behaviors in different fields or the many possible database errors—have also started to perform analyses of citations for a number of purposes. To be mentioned are:

1. Ranking of journals, often stimulated by decreasing library budgets and the wish to keep the “best” (i.e. most cited in two years) journals. Also, senior scientists and their PhD students and postdocs use these rankings to decide to which journal they will submit their papers. Editors and editorial staff members do have an interest in these rankings and to climb on the ladder (see below), even though they realize that IF ranges vary per science area.
2. Ranking of scientists; find out who has most citations, or most citations

[\*] The two-year IF of a journal is defined as the number of citations in year *X* to all papers published by the journal in years *X*–1 and *X*–2, divided by the number of citable papers (as defined by ISI) in that journal in the years *X*–1 and *X*–2.

[\*\*] Most readers are aware of different citation rates in different fields, which naturally depend on the number of researchers; intensely investigated fields are expected to generate more citations per published paper and higher IF for journals in these fields. Two new Elsevier parameters, SNIP and SJR, correct for these: [www.scimagojr.com](http://www.scimagojr.com)

[\*\*\*] The role (and power) of ISI/Thomson Reuters should not be underestimated in rankings. For example, their criteria about what is a “citable paper” are not public.

[\*] The *h* index is determined as the number of papers (*h*) from a person (or group) in a certain period (normally a whole career), that has been cited at least *h* times.

[\*] Prof. Dr. J. Reedijk  
Leiden University (The Netherlands)  
and  
King Saud University, Riyadh (Saudi Arabia)  
E-mail: [reedijk@chem.leidenuniv.nl](mailto:reedijk@chem.leidenuniv.nl)

per paper; and more recently the value of the *h* index, for the whole career, or part of it if wished.

3. Rankings of groups, departments, institutes and even universities, for example, in the well-known Shanghai rankings, where citations form a major component, and whole countries.

So one can observe strong driving forces to improve such numbers, in some cases even with unfair means. Funding, salaries or bonuses of staff, and attractiveness to (foreign) students increasingly depend on rankings, and these in turn are determined by, or at least include, citations to research papers and parameters derived from them. So, it is a challenge to explore and apply methods to obtain more citations: to a paper, to a journal, to a scientist. Consequently, optimizing conditions for IFs are available, and unfortunately, fabricating or engineering of IFs and citations in general has also been observed and will probably increase.

### Journal Editors

Journals should make sure that citations are correct. An editorial policy to increase the impact factor has been the introduction of certain categories of papers that might result in above-average citations, such as different sorts of reviews. As many journals do this, the effect appears to be marginal after a while—just a small increase in IF for all. Editors, of course, should primarily go for quality and should be allowed to hunt for top quality and hot material, likely to be read—and cited. Perhaps this explains why we nowadays see so many editorial staff at scientific meetings.

Journals will tend to accept papers that are from “fashionable” research areas (and of good quality) and to give lower priority of acceptance to papers that are within the scope of the journal, high quality, but less fashionable (read: less likely to be cited). This behavior can be defended when the number of acceptable manuscripts submitted is very much

larger than the annual volume can handle.

Editors nowadays can easily put potential hot papers “on hold”. For example, an article might be ready for publication in mid-2011 but only online (with a DOI number). If page numbers for a 2012 issue are added later, the effective reading and citation time window increases significantly, since ISI will count the article as being from 2012. Here I see a potential unethical temptation.

Highlights in other periodicals of the same publishing house about recent important or successful papers in their journals are also used. Of course, this can be an excellent service to the readers of the journals, but the mentioned citations will count towards the IF. It is, however, unethical when an editorial team writes a paper (for five other periodicals) praising and citing large numbers of hot papers in their journal. It is also unethical to have in almost every issue of the journal an editorial (non-citable item) that, perhaps exclusively, cites several papers from the last two years published in that particular journal. As a “service to the reader” some very good papers in the same journal are mentioned, discussed, and cited. For 12 non-citable editorials each year with 5–10 same-journal citations each, a total of say 300 citable publications each year, these editorials add 0.2–0.4 units to the two-year IF.

Some editors also routinely encourage, or even require (more) citation(s) to the journal. So upon acceptance of every paper, the editor may require an additional two references to the journal from the previous two years, thereby adding 2.0 to the two-year IF! Such journals may be suspended from JCR for some time, as has happened in several cases; see: <http://wp.me/pcvbl-5K0>. More details of these and other editorial policies to engineer IFs are discussed elsewhere.<sup>[2]</sup>

### Scientists and University Administrators

To improve status and ranking, universities worldwide are often successful in

hiring highly ranked scientists from other places. Well known in the UK are cases of staff recruiting before a certain “transfer date”; these transfers have been common practice before national university research evaluations.

Common and relevant issues for individual scientists include making sure that a scientist and her or his work is correctly and completely used in evaluations. It can be problematic when scientists change their name, or the spelling of it, or even their initials, during their career. Also, universities and journals have learned to be careful in changing names and postal addresses.

A well-known practice is to send one’s own recently published paper to many colleagues, stressing its importance, in the hope they will cite it. Citing disproportionately more of your own, earlier, papers no doubt happens on a large scale, and it also occurred before the ranking craze; self-citations, however, can be easily eliminated in an analysis. Scientists who worked and co-authored highly cited work as PhD students or postdocs will, in any case, benefit from this association their entire career. One cannot do much later to increase his or her own personal *h* index, other than publish good papers and wait.

### Holding Up Standards

Even if the use of citation analysis in research assessment eventually would decrease, as searching the literature and citing will be done more based on databases and less by reputation only, bibliometrics, counting, rankings, and evaluations will remain for a while. Peer-reviewed scientific publications are the basis of scientific evaluations, and they must adhere to the highest ethical standards. These standards should be the same for all authors, referees, and editors! Professors also have a special duty to teach correct ethical citation behavior to their PhD students, and postdocs. Journal impact factors may be affected by particular editorial strategies, be it intentionally or not. Therefore, one must be most careful in interpreting and using journal impact factors; authors, editors, and policy makers must be

aware of their potential manipulability. Also values of the  $h$  index for individuals must be dealt with carefully, not so much for manipulation as for differ-

ences in research area and whether or not a researcher has co-authored highly cited articles during postdoc or PhD work.

[1] E. Garfield, *Science* **1972**, *178*, 471-479.

[2] H. F. Moed, J. Reedijk, *J. Doc.* **2008**, *64*, 183-192.

**ChemistryViews**

Spot your favorite content  
[www.ChemistryViews.org](http://www.ChemistryViews.org)

Education & entertainment

Exciting news

Unique articles

Multi-media

Free & easy access to new magazine

New online service brought to you by

ChemPubSoc Europe

WILEY-VCH

603701008