Research evaluation

Citation versus expert opinions

Research evaluation judgements appear to be moving away from expert opinions, with more emphasis instead resting on citation counts. Professor Lawrence Smolinsky and his colleagues from Louisiana State University, USA, examine whether peer reviews and citation metrics measure similar notions of an article's importance, impact, and value. They explore the relationship between the various phenomena related to the perceived importance of subfields when measured by citations and experts. In addition, they provide an innovative theoretical framework to measure the consequences of bibliometric methodology.

Research evaluation which attempts to measure the quality of scientific articles has a significant impact on judgements concerning research being carried out by scientific institutions, universities, individual scientists, and research groups. Previously, these research evaluations factored in expert opinions. Now, however, such judgements instead commonly rely on citation counts or indicators based upon them. For example, The National Research Council, the operating arm of the National Academies of Sciences, Engineering, and Medicine, carries out evaluations of graduate departments in the US. In 2010, the rankings put emphasis on data including citation counts, whereas previous rankings stressed the importance of the surveys of scholars. Likewise, a research journal's ranking is now measured almost entirely by citation count data.

Professor Lawrence Smolinsky, from the Department of Mathematics at Louisiana State University, questions whether peer reviews and citation metrics measure similar notions of an article's importance, impact, and value. Together with Professor Daniel Sage, Dr Aaron Lercher, and Aaron Cao, Professor Smolinsky examines how citation metrics and peer review processes are used to discern the value of scientific papers.
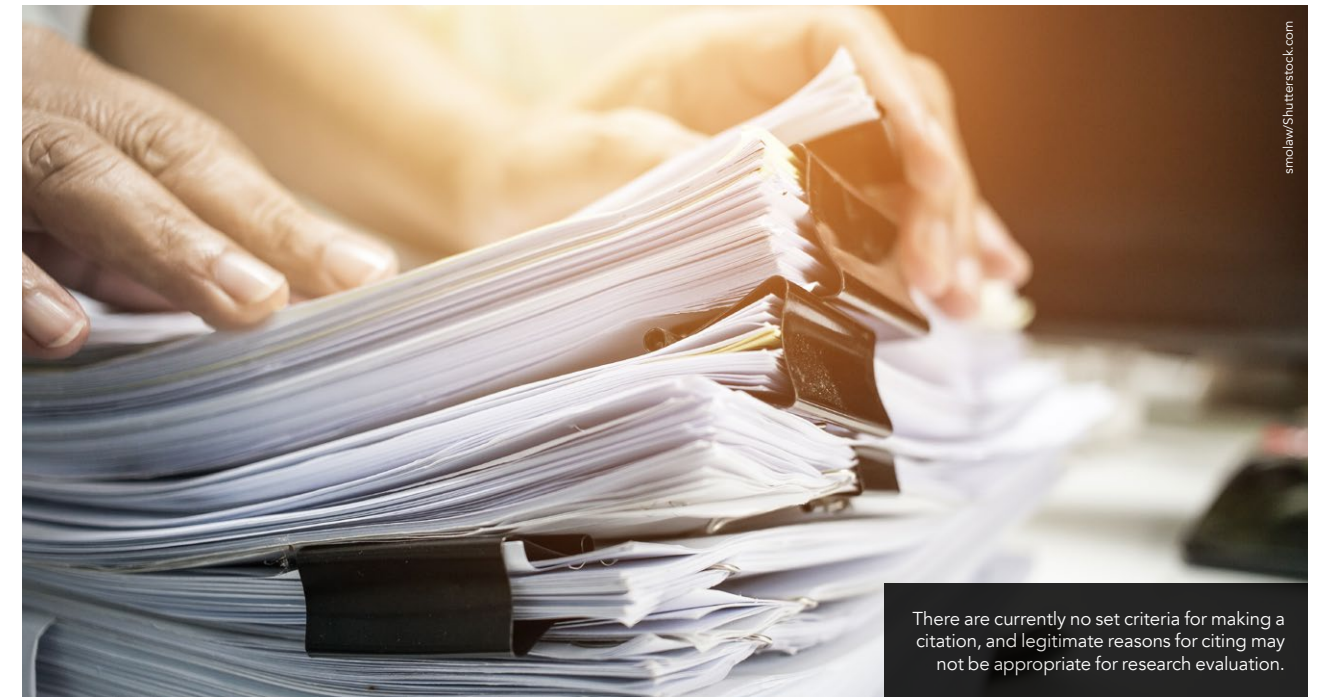
PEER REVIEW

Peer review is used to directly assess the value, quality, and relevance of a broad array of scholarly work. It includes reviews of individual articles submitted for publication by referees and editors, reviewing proposals for grant support, and reviewing for scholarly prizes, awards, and honours. Peer review

aurielaki/Shutterstock.com

is therefore a serious professional responsibility. The reviewer is serving as an independent expert, and it is a matter of professional ethics to remain impartial. In cases where a significant conflict of interest exists, subject experts are expected to recuse themselves in the interests of fairness and rigorous assessment. It is important to note, however, that the methodology employed by peer reviewers is not uniform. For instance, different journals can give diverse instructions to their reviewers, and they can also use the information in a variety of ways. There is also potential for personal bias as some journals and funding organisations allow researchers to suggest and exclude potential reviewers. Further, the lack of organised systematic and publicly available peer review data can make examining this method difficult.

CITATION ANALYSIS

Citation counts of research publications are widely used within information science as an instrument of research evaluation, yet there are no set criteria for making a citation. It is often assumed that citations are the correct metric of quality or impact, but that assumption is fundamentally based on speculation as to why some articles are chosen to be cited and others are not. In the absence of verified criteria for the citation process, the weight placed upon the metric of citation as an indicator of value needs to be closely scrutinised rather than simply accepted. Rather than a measure of quality, citation data is sometimes restricted to being called a measure of scientific impact. However, even this restriction requires scrutiny except in the logically circular approach of simply defining the scientific impact of an article in terms of its citations.

Smolinsky explains how a YouTube video viewer can select 'like' because they truly like a video, and without any

other motivation. A citation, however, is not a 'like', and an author can have a variety of motives to include a citation. An author may cite an article because the citation is essential, but they also want to promote their article and may make citations accordingly. Smolinsky believes it is a legitimate reason for authors to cite articles but may be confounding for research evaluation. There are also issues arising from negative citations, self-citations, methodological and review articles, journal prestige, and the variations that arise by discipline. There is no simple way to encapsulate the meaning of citations as the citer is not anonymous, thus the reference may be made from self-interest. Furthermore, the only awarding standard for the citation is its relevance or desire in the eye of the author, leaving their use in evaluation unclear.

MATHEMATICS AS A DATA SOURCE

Smolinsky and his colleagues selected mathematics as a laboratory for investigating the consistency between citations and peer review. Since mathematics is an exact science there is a narrow range of reasons for citations. It also has a lower average number of joint authors per article than other sciences, and as there is no laboratory

There are currently no set criteria for making a citation, and legitimate reasons for citing may not be appropriate for research evaluation.

work, there are fewer collaborator and fewer 'team self-citations.'

The researchers used Mathematical Reviews (MR) and the Web of Science (WOS) as data sources for their informetrics (the study of quantitative aspects of information) research. From 1993 to 2004, MR selected particularly

> *An author may cite an article because the citation is essential, but they also want to promote their article and may make citations accordingly.*

significant articles for featured reviews in MR. Of these elite featured review articles, 734 also had citation counts in the WOS. During the same period, 1,559 WOS articles with mathematics in their categories were highly cited with more than 100 citations and also listed in MR.

Examining the bibliometrics (meaning the quantitative publication and citation data used to measure research impact) of these articles, the researchers note that these highly cited and featured review articles are not evenly spread throughout the various subfields of mathematics, and the distribution reflects the perceived importance of the subfields. The hiring patterns in top mathematics departments and the mathematician's interests are also

more closely related to this perceived importance of subfields than to citations.

CITATION ANALYSIS VERSUS PEER REVIEW

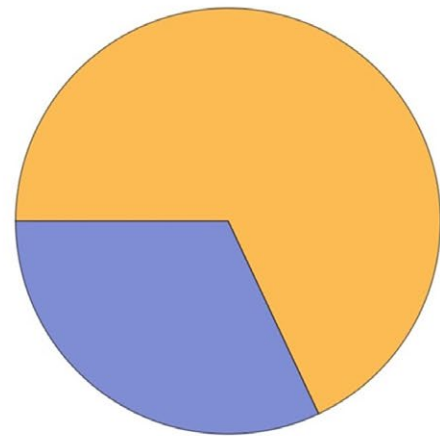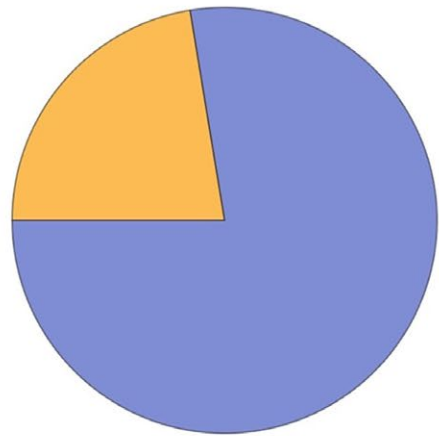The researchers investigated to what extent the measure of value obtained using citations is similar to the measure of value obtained using peer reviews. They had two main aims. Firstly, to find out if there was a statistically significant correlation between the citations and peer opinion. Secondly, they wanted to uncover if citations and peer review measure a common notion.

FEATURED REVIEW ARTICLES VERSUS HIGHLY CITED ARTICLES

The researchers found that only 7.83% of the highly cited WOS mathematics articles were featured MR articles. Similarly, only 16.62% of the featured MR articles were highly cited on the WOS. This suggests negligible relationships between peer selection and being highly cited. The analysis also revealed that being described as a featured review article and being highly cited are markedly distinct, demonstrating that peer review and citation counts tend to provide different determinations of highly distinguished articles.

smolaw/Shutterstock.com

Publications in physics review journals, in the subfields of physics of elementary particles, field theory, gravitation, and cosmology, as a share of all physics publications. The left shows the actual proportion of publications and the right shows the proportion of credit given using total author counting. This illustrates the stark difference that total author counting may cause.

Legend:
- Elementary particles, field theory, gravitation, & cosmology
- All other subfields of physics

SUBFIELD FINDINGS

Exploring the relationship between the various phenomena related to the perceived importance of subfields when measured by citations and experts uncovered a very strong correlation between hiring in the top departments and the featured review articles. The correlation between hiring and the highly cited articles was still strong, but less so. Faculty interest had a stronger correlation with the featured review article than either the highly cited articles or hiring subfields.

A DISCONNECT BETWEEN PEER REVIEWS AND CITATION COUNTS

These findings, together with the current trend of viewing citation counts as the primary metric for the value and impact of an article, point to a disconnect between high peer measures and high citation counts. Smolinsky and his colleagues remark that this 'may amount to a shift in the very meaning of value and impact used in describing academic articles.' This was illustrated in the 2010 National Research Council's assessment of US research doctorate programmes where, rather than employing direct expert opinion to assess the quality of research faculty, it was only used to weigh the importance of various data sets. Experts could express the relative weight of citations per publication, but not offer an opinion as to whether this was an adequate measurement of quality. This was further confounded with the use of co-author weighting, and partly motivated another study by Smolinsky and Aaron Lercher.

CO-AUTHOR WEIGHTING

The researchers found that 'there is a substantial difference in the value or impact of a specific subfield depending on the credit system employed to credit individual authors.' The impact is here measured by conventional citation and article counts. This is likely to have consequences both for institutions and for whole disciplines.

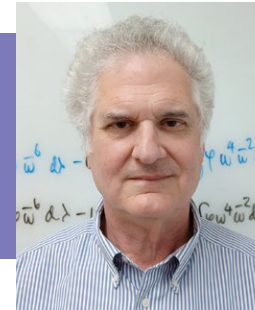A common way to credit citations and articles to authors is *total author counting*. This gives each author of an article full credit for the full citation count and the article. For example, if an article has 100 citations and 10 authors, then each author is given credit for 100 citations and for authoring one article. This method is generous to authors and inadvertently values research areas where articles have large numbers of co-authors above those fields with small numbers of co-authors.

To demonstrate their innovative theoretical framework used to measure the consequences of bibliometric methodology, the researchers carried out a study of almost 15,000 physics articles published in the American Physical Society's *Physical Review* journals from 2013 to 2019. They assigned subdiscipline classifications to articles and gathered citation, publication, and author information. They introduced *citation credit spaces* to examine the total amount of credit given to an area. Considering the previous example, a paper with 100 citations and 10 authors gets credit for 1000 citations. The researchers call this *co-author weighted citation credit space*. Contrastingly, an article with 100 citations but only one author, is only awarded a co-author credit space of 100 citations.

PRACTICAL IMPLICATIONS

Smolinsky and his colleagues have shown that the selection criteria for expert opinions and citation counts of articles are largely different. Measurement by citation counts is cost efficient as the work is being done commercially for reasons unrelated to evaluation, making it inexpensive and available. Smolinsky warns, however, that 'the method of crediting individual scientists has consequences beyond the individual and affects the perceived impact of whole subfields and institutions'.

> *Smolinsky and his colleagues selected mathematics as a laboratory for investigating the consistency between citations and peer review.*

Frannyanne/Shutterstock.com

Behind the Research
Professor Lawrence Smolinsky

E: smolinsky@math.lsu.edu   T: +1 225 456 3610

Research Objectives

Professor Lawrence Smolinsky and his colleagues examine the discrepancies between peer review and citation metrics as reliable indicators of research quality.

Detail

Address
Department of Mathematics, Louisiana State University, Baton Rouge, LA USA, 70803–4918

Collaborators
Aaron J Lercher, Daniel S Sage, Aaron Cao

References

Smolinsky, L, Sage, DS, Lercher, AJ, Cao, A, (2021) Citations versus expert opinions: citation analysis of featured reviews of the American Mathematical Society. *Scientometrics*, 126(5) 3853–3870. Pre-print version: doi.org/10.48550/arXiv.2005.05389

Smolinsky, L, Lercher, A, (2020) Co-author weighting in bibliometric methodology and subfields of a scientific discipline. *Journal of Data and Information Science*, 5(3), 84–96. doi.org/10.2478/jdis-2020-0021

Personal Response

What sparked your interest in research evaluation?

My first foray into the field was during my service as department chair and was for practical reasons. I had proposed relocating our department to the College of Science. Faculty members were concerned about comparisons in funding and bibliometrics with the experimental sciences and between subfields in mathematics. I began investing the differences to clarify the standards for our potential new dean and for the faculty. That investigation and the subsequent memos to the dean on cross-discipline and subdiscipline comparisons contributed to my first papers in information science. My interest and research grew in the following years.

LSU
LOUISIANA STATE UNIVERSITY