RESEARCH PAPER

Interdisciplinary Comparison of Scientific Impact of Publications Using the Citation-Ratio

Arthur R. Bos[1] and Sandrine Nitza[2]

[1] Department of Biology, American University in Cairo, New Cairo 11835, EG

[2] University of Abertay, Dundee, Scotland DD1 1HG, UK

Corresponding author: Arthur R. Bos (arbos@aucegypt.edu)

The commonly used indexes for evaluating the scientific impact of publications and individual researchers do not allow accurate comparison between disciplines with varying citation frequencies. The Citation-Ratio (CR) was developed to measure impact of an individual publication and allow field-normalised comparison. The CR equals the total number of citations of a publication divided by the median of citations of its references and was tested for the top 5% of the most-cited publications of 13 selected disciplines in sciences, social sciences and humanities. Each publication had a CR = 0 until it was firstly cited. At CR = 1 the number of citations equalled the median of citations of the references. CRs of the most-cited publications mostly ranged between 1 and 10 and were not significantly different across the selected disciplines. In contrast, the total number of citations of the same publications were significantly different across disciplines. One of the advantages of the CR is that it can be calculated for any publication as long as it has references (e.g. books, book chapters, reports, and symposium contributions).

Keywords: Academic discipline; Citation index; *h*-index; Journal impact factor; Peer review

Introduction

Borders between disciplines were relatively well defined in the distant past, but have faded with the increasing development of the sciences. Today, interdisciplinary research is highly valued, and scholars are encouraged to collaborate with colleagues from other research areas (Porter & Rafols 2009). Furthermore, a scientist's research may shift from one discipline to another in the course of their careers. Consequently, scientists may publish in various disciplines and produce multiple publications with incomparable citation patterns (Wagner et al. 2011). Therefore, we hypothesised that each publication has its own niche within research; its own set of referenced publications and therefore, a unique group of peer scientists. Hereby, each publication is an entity in itself with its own individual community of interest and thus would not be primarily related to the journal it was published in.

It has been common practice that the impact of a publication is simply summarised by using the Journal Impact Factor (JIF) without considering its actual number of citations. This measure of impact is highly inaccurate, as the JIF equals the mean citation frequency of all papers published in a journal in one year and thus does not correctly represent all papers. Moreover, the use of the arithmetic mean is mathematically inappropriate, because distributions of citations are almost always skewed (Bornmann et al. 2008). Therefore, in many cases the JIF under- or over-estimates the exact number of citations of a publication. Moreover, JIFs highly vary between academic disciplines (Castellano & Radicchi 2009).

From the above, we conclude that another approach is needed to accurately summarise the impact of an individual publication, let alone the collective work of an author or a group of researchers. Schubert & Braun (1993) proposed an indicator for evaluating journal impact based on the relationship between the mean number of citations of publications and the expected mean number of citations of the journal in question. This approach approximates our intended design for evaluating individual publications, but adjustments are needed. As mentioned above, the arithmetic mean does not correctly represent the non-parametric

distributions of publications' citations (Bornmann et al. 2008) and an alternative central tendency must be selected. Furthermore, we postulate that the quality of a publication is determined by the utilisation (=number of citations) of its unique group of peer researchers. Its citations should therefore be compared to the work's most relevant literature, which is the set of references it has used, and not to the performance of the journal.

The present paper aims to develop a measure that evaluates the impact of a publication independent of the citation frequency across multiple disciplines.

Methods

Thirteen academic disciplines were selected from university departments covering an as broad as possible academic spectrum (sciences, social sciences, and humanities; **Table 1**). Each discipline keyword was used to conduct a topic search in the Web of Science (Thomson Reuters 2016) in December 2015 and January 2016. The top 5% of most-cited publications of the year 2000 were systematically sampled ($n = 10$). As the number of top publications varied per discipline, a step size of sampling was determined by dividing the discipline-dependent number by 10 (**Table 1**). The analysis started with the most frequently cited publication, moving on to the next, based on the step size of each discipline. This method was continued until 10 publications were sampled per discipline. The JIF-2014 and total number of citations of each publication were recorded. Subsequently, the number of citations for each reference in one publication were recorded. If a reference had more than 60 citations, then only the first 60 were included in the analysis (**Table 1**). The vast majority of selected publications were regular research papers, while six disciplines (Astrophysics, Chemistry, Education, Immunology, Meteorology, and Psychology) were also represented by one or two review papers. A publication in which authors responded to a published article, hence one reference, was included for Philosophy (**Table 1**).

The Citation-Ratio (CR) was calculated for each publication by dividing the total number of citations by the median of the total citations of all references. It was decided that the median would be used instead of the arithmetic mean in order to avoid individual publications impacting the denominator of the CR too highly. The total number of citations and CR were tested for differences between disciplines using ANOVA after log transformation, assuring variance equality between the data sets (Levene's-test, $P > 0.05$). Tukey's honestly significant difference (HSD) test was used for *post-hoc* analysis ($\alpha = 0.05$).

Table 1: Total number of publications from the year 2000 with sampling step size for the 5% most-cited publications and the median number and range of references per publication for thirteen disciplines searched in the Web of Science.

Discipline	Publications	Sampling step size	References per Publication	
	Total number	**of top 5%**	**Median**	**Range**
Agriculture	2,458	10	44.0	19–117
Astrophysics	260	1	91.0	21–606
Archaeology	697	3	53.0	18–945
Chemistry	72,612	350	52.5	28–139
Education	26,648	100	61.0	32–114
Genetics	93,398	500	59.5	23–166
Immunology	34,008	150	43.5	19–291
Music	2,572	10	50.5	13–269
Neurology	853	4	45.5	12–188
Meteorology	281	1	54.0	18–220
Nanotechnology	173	1	32.0	9–435
Philosophy	3,197	15	57.5	1–163
Psychology	26,836	100	63.5	19–149

Results

Citation frequency of the top 5% of the most-cited publications was significantly different (ANOVA, $F_{12,117}$ = 3.59, $P < 0.01$) between 13 selected academic fields (**Figure 1a**). *Post-hoc* analysis revealed that citation frequency of philosophy was significantly lower than of chemistry, genetics, immunology, and nanotechnology (HSD-test, $P < 0.05$). Furthermore, citation frequency of music was significantly lower (HSD-test, $P < 0.05$) than of chemistry, genetics, and nanotechnology (**Figure 1a**). The highest total number of citations of a single publication was 14,444 in Chemistry, whereas the lowest total number of citations of a publication was 37 in Philosophy.

The CR equals zero until a publication receives its first citation and may gradually increase when the publication is continously cited. At a value 1, the CR reaches a balance between its own citations and the median of citations of all references. The vast majority of the top 5% of the most-cited publications had a CR between 1 and 10 across all disciplines (**Figure 1b**). The highest CR was 137.6 for the most-cited publication in chemistry. The CR of the top 5% of the most-cited publications was not significantly different between academic disciplines (ANOVA, $F_{12,117}$ = 1.27, $P = 0.244$).

Discussion and Conclusions

Academic disciplines have varying citation frequencies, which has made comparison of scientific impact of publications and authors across disciplines difficult (e.g. Lillquist & Green 2010). As a result, impact indexes have been biased toward disciplines with high citation frequencies. The CR is a measure that does not discriminate between disciplines by compensating for citation frequency differences. The CR, instead, detects
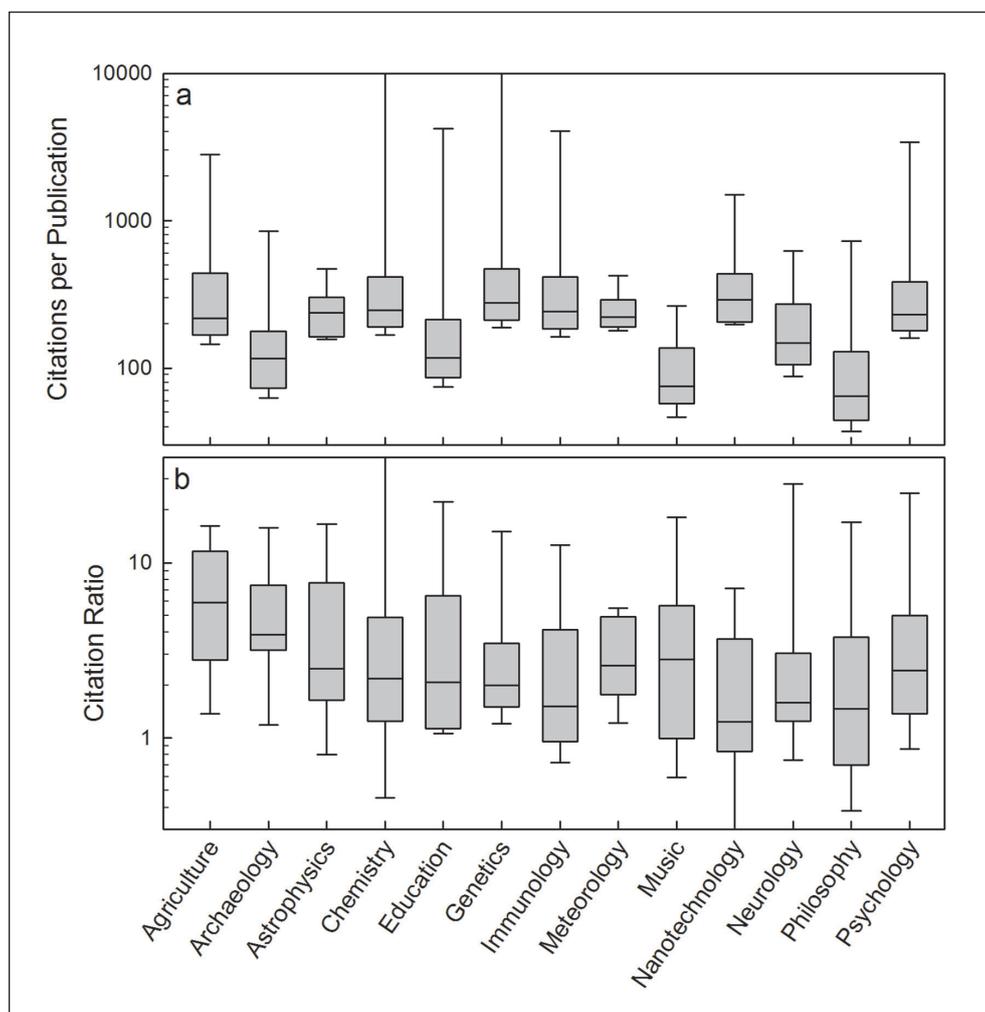

Figure 1: Total citations per publication **(a)** and Citation-Ratio **(b)** of the top 5% of the most-cited papers published in 2000 for thirteen academic disciplines.

publications with an above-average number of citations in their own specialised research niche. As well as the advantage of being citation frequency independent, the CR can be calculated for any publication that uses references. The CR is thus not limited to journal articles, as is the case for indexes using JIFs (Glänzel & Moed 2002), but instead, can also be determined for books, book chapters, technical reports, symposium papers, etc. Calculating impact of information sources without references, such as websites and blogs, is out of scope of the CR.

All impact measures have their own advantages and disadvantages (e.g. Egghe 2010). Therefore, it is important to ensure that any measures aiming to evaluate scientific impact cannot be manipulated by those which are being evaluated. To analyse the likelihood of manipulation of the CR, the role of the equations' numerator (number of citations) and denominator (median of citations of references) will be discussed separately. The main concern with the numerator is that it can be affected by self-citations, similar to reports for the h-index (Hirsch 2005; Bartneck & Kokkelmans 2011). When using the CR, a self-citation affects two publications: the citing publication as well as the cited publication. After self-citation, the numerator of the cited publication increases by one, while the denominator of the citing publication may increase causing the overall CR to decrease. Since the value of the denominator is determined by a median however, a single citation of one reference may have a limited effect. Although a self-citation may thus have a larger effect on the numerator than on the denominator, self-citations will not easily lift the CR. Excluding self-citations from calculations could be considered an option, since authors could no longer influence CRs. However, past attempts of doing so for other indexes have shown limited impact by self-citations (Thijs & Glänzel 2006) and have proven to be difficult due to ambiguity regarding article assignment to authors with commonly used names (Han, Zha and Giles 2005).

While not accounting for the limited effect of self-citations, there are mainly two options to manipulate the CR's denominator. First of all, an author could avoid using references with many citations, and secondly, an author could choose to increase the number of references by selecting unnecessary publications with few citations or limited citation potential.

The avoidance of highly cited papers seems to be a strategy that leaves out reviews and standard works, which highly determines the unique citation frequency of a particular discipline. Leaving out such references could possibly result in a publication not being recognised as a contribution to its field and, as a consequence, it may not be cited as often. On the other hand, extending a reference list by adding publications with limited citation potential would create an unnecessarily long list. Regular journal contributions generally include 10 to 60 references, whereas reviews can refer to up to several hundreds of publications. Although disciplines may have different practices in selecting the total number of references, the future development of a publication's CR is difficult to forecast for contributions with many references. However, there are a few contribution types, such as one-page notes and short communications that may have as little as three references (e.g. Bos, Gumanao and Salac 2008). Manipulation by self-citations may therefore be easier in such contributions. The disadvantage of short contributions is that they contain very limited information and are consequently, less frequently cited (Vieira & Gomes 2010). Moreover, citation frequency of short contributions containing unique or new knowledge may only be high for a finite period. Once the novelty of the contribution fades, the citation frequency may rapidly decline, which could then lead to a decreasing CR. This exemplifies that the CR changes with every new citation to the paper in question as well as to its references. Therefore, if a publication has reached its citation peak and its references continue to be cited at their same frequency, its CR will slowly decrease, meaning that the contribution's impact gradually diminishes. In contrast, the nature of the h-index causes it only to increase (Hirsch 2005), as citations are never withdrawn.

For both options of manipulating the CR's denominator, there is an important role for journal editors and reviewers, who must assure that standard works are included and that reference lists have reasonable lengths. When encouraging authors to select a low number of references, they must select publications with highest relevance to the subject, without considering the number of citations. Apart from these potential risks of slightly increasing the CR, it would be difficult to forecast the citation success of one's own publication and, more importantly, of recent work published by others. Given the patterns discussed above and the important role of editors and reviewers, we believe the manipulation potential of the CR is low.

We conclude that the CR is a useful and promising tool for comparing scientific impact of publications across disciplines and potentially for interdisciplinary works and suggest that the CR's application is further tested on large databases of an extended set of disciplines.

Competing Interests

The authors have no competing interests to declare.

References

Bartneck, C and **Kokkelmans, S.** 2011. Detecting h-index manipulation through self-citation analysis. *Scientometrics*, 87(1): 85–98. DOI: https://doi.org/10.1007/s11192-010-0306-5

Bornmann, L, Mutz, R, Neuhaus, C and **Daniel, HD.** 2008. Use of citation counts for research evaluation: Standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in Science and Environmental Politics*, 8: 93–102. DOI: https://doi.org/10.3354/esep00084

Bos, AR, Gumanao, GS and **Salac, FN.** 2008. A newly discovered predator of the crown-of-thorns starfish. *Coral Reefs*, 27(3): 581. DOI: https://doi.org/10.1007/s00338-008-0364-9

Castellano, C and **Radicchi, F.** 2009. On the fairness of using relative indicators for comparing citation performance in different disciplines. *Archivum Immunologiae et Therapia Experimentalis*, 57: 85–90. DOI: https://doi.org/10.1007/s00005-009-0014-0

Egghe, L. 2010. The Hirsch-index and related impact measures. *Annual Review of Information Science and Technology*, 44: 65–144. DOI: https://doi.org/10.1002/aris.2010.1440440109

Glänzel, W and **Moed, H.** 2002. Journal impact measures in bibliometric research. *Scientometrics*, 53(2): 171–193. DOI: https://doi.org/10.1023/A:1014848323806

Han, H, Zha, H and **Giles, CL.** 2005. Name disambiguation in author citations using a k-way spectral clustering method. In: Proceedings of the 5th ACM/IEEE-CS Joint Conference, Denver, TX, 334–343. DOI: https://doi.org/10.1145/1065385.1065462

Hirsch, JE. 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46): 16569–16572. DOI: https://doi.org/10.1073/pnas.0507655102

Lillquist, E and **Green, S.** 2010. The discipline dependence of citation statistics. *Scientometrics*, 84(3): 749–762. DOI: https://doi.org/10.1007/s11192-010-0162-3

Porter, AL and **Rafols, I.** 2009. Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics*, 81(3): 719–745. DOI: https://doi.org/10.1007/s11192-008-2197-2

Schubert, A and **Braun, T.** 1993. Reference standards for citation based assessments. *Scientometrics*, 26(1): 21–35. DOI: https://doi.org/10.1007/BF02016790

Thijs, B and **Glänzel, W.** 2006. The influence of author self-citations on bibliometric meso-indicators. The case of European universities. *Scientometrics*, 66(1): 71–80. DOI: https://doi.org/10.1007/s11192-006-0006-3

Thomson Reuters. 2016. *Web of Science.* Available at http://webofknowledge.com [Last accessed 14 January 2016].

Vieira, ES and **Gomes, JANF.** 2010. Citations to scientific articles: Its distribution and dependence on the article features. *Journal of Informetrics*, 4: 1–13. DOI: https://doi.org/10.1016/j.joi.2009.06.002

Wagner, CS, Roessner, JD, Bobb, K, Thompson Klein, J, Boyack, KW, Keyton, J, Rafols, I and **Börner, K.** 2011. Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature. *Journal of Informetrics*, 165: 14–26. DOI: https://doi.org/10.1016/j.joi.2010.06.004