



Deep and narrow impact: introducing location filtered citation counting

Dangzhi Zhao¹ · Andreas Strotmann²

Received: 26 May 2019 / Published online: 15 November 2019
© Akadémiai Kiadó, Budapest, Hungary 2019

Abstract

The present study tests a citation counting method that filters out citations in the introductory and backgrounds sections and then weighs the remaining citations by their in-text frequency. The dataset used comprises articles on bibliometrics available in full text in PubMed Central. This method was inspired by findings from previous studies that in-text frequency indicates importance of citations and citations in Methodology, Results, Discussion, and Conclusions sections tend to be more important to a citing article. We found that this method makes a large difference in author ranking as suggested by a 0.4 correlation between ranking by this method and that by traditional citation counting. Generally, this method has ranked authors concerning biomedical issues higher and those focused on bibliometrics or science communication issues lower compared to traditional citation counting. This rank change pattern suggests that this method appears to have made essential citations stand out more, i.e., citations that studies concerning biomedicine are expected to draw on more heavily. This method has also ranked guidelines or theoretical or methodological frameworks for systematic reviews, meta-analyses, knowledge translation, and scoping studies much higher, indicating that Bibliometrics has been mostly employed in these types of studies in biomedical fields. Unfortunately, citation network analysis doesn't seem to have been employed much as indicated by key authors representing science mapping being ranked much lower by this method although it has been shown to be informative for these types of studies.

Keywords Citation analysis · Weighted citation analysis · Location-filtered citation counting · Research evaluation · Bibliometrics · Biomedical research fields

Introduction

Citation analysis is used in research evaluation exercises around the globe, directly affecting the work and lives of millions of researchers and the expenditure of billions of dollars. It is therefore crucial to address any problems or limitations that plague it. Central amongst

✉ Dangzhi Zhao
dzhao@ualberta.ca

¹ University of Alberta, Edmonton, AB T6G 2J4, Canada

² ScienceXplore, F.-G.-Keller-Str. 10, 01814 Bad Schandau, Germany

critiques of practices of citation analysis has long been that it treats all citations equally, be they crucial to the citing paper or perfunctory. This problem is especially troublesome when tracing or assessing research impact. Weighing citations by how they are used in the citing paper has therefore long been proposed as a solution to this problem (Herlach 1978; Narin 1976; Voos and Dagaev 1976). By weighing citations, it is hoped that essential citations could be assigned greater weight than perfunctory ones so that citation analysis can focus on more profound influences and organic relationships, and research evaluation, knowledge network analysis, knowledge representation, and information retrieval aided by citation analysis can be improved (Zhao et al. 2017; Zhao and Strotmann 2015b, 2016).

Studies have consistently found that in-text frequency of a cited reference indicates its importance (Bonzi 1982; Chubin and Moitra 1975; Herlach 1978; Tang and Safer 2008; Voos and Dagaev 1976; Zhao et al. 2017; Zhu et al. 2015). Based on the assumption that the more frequently a reference is mentioned in the text, the more significant it is to the citing paper, frequency-weighted citation analysis approaches assign a weight of N (or a function of N such as N^2) to a citation that appears N times in a citing paper (Bu et al. 2018; Zhu et al. 2015). However, Zhao et al. (2017) found that quite a large percentage of multi-citations appear to play purely a nonessential role in the citing paper, and could be overweighted by frequency-weighted citation counting. This finding underscores the importance of filtering out nonessential citations before assigning weight in order to improve the accuracy and effectiveness of frequency-weighted citation analysis. Future studies were invited to explore effective ways to filter out nonessential citations, and to evaluate the differences that filtering out nonessential citations before assigning weight might make in weighted citation analysis. The present study is such an attempt. It explores how much of a difference it might make in research evaluation to filter citations by their in-text location. It also examines patterns of bibliometrics-related studies in the biomedical research fields where both wide applications of and serious concerns about citation analysis for research evaluation have been present.

Background and research questions

Although weighing citations by how they are used in the citing paper has long been proposed as a theoretical solution to the problem of equal treatment of perfunctory, significant or crucial citations, in practice, it has not been studied closely at a large-scale until recently. Increasingly available digital full-text documents and advances in text processing technologies are now making it feasible to conduct large-scale studies on weighted citation analysis. As a result, this type of studies has attracted increasing research interest in recent years. Studies have experimented with weighing citations by the frequency with which they occur in the text (e.g., Ding et al. 2013; Hou et al. 2011; Tang and Safer 2008; Zhu et al. 2015), by the citation impact of citing papers (Ding and Cronin 2011), and by the location and context in which they are cited (Boyack et al. 2013; Jeong et al. 2014). Zhu et al. (2015) found that in-text citation frequency was the best of many full-text features (including citation location) to help spot citations that were considered crucial to the citing papers by their authors. Studies have also examined characteristics of in-text citations across research fields, especially their distribution within the full-text (Bertin et al. 2016; Boyack et al. 2018; Hu et al. 2017; Hsiao and Chen 2018; Pak et al. 2018; Otto et al. 2019; Thelwall 2019a, b). Datasets specifically constructed to support in-text citation analysis studies using natural language processing have also begun to emerge (Bertin and Atanassova 2018).

The basic assumption underlying citation analysis is that a citation represents the citing author's use of the cited work (Garfield 1979; White 1990; Zhao and Strotmann 2015a). However, scholars do cite for various reasons and citations do serve many different functions in citing papers. Beginning in the 1970s, a great deal of research has been done on citer motives, citing behaviours, and citation functions. Small (1982), for example, identified five typical distinctions in citation classification schemes: (1) negative or refuted, (2) perfunctory or noted only, (3) compared or reviewed, (4) used or applied, and (5) substantiated or supported by the citing work. It was also around this time that the use of citation analysis in research evaluation caused concerns that a citation may not represent actual use of a cited document, and that citation counts that do not take into account citers' motives, citing behavior, and citation functions may not reflect the impact or merit of the cited documents (Brooks 1985, 1986; Case and Higgins 2000; Chubin and Moitra 1975; Garfield 1979; Liu 1993; Moravcsik and Murugesan 1975; Shadish et al. 1995; Vinkler 1987; White and Wang 1997). These studies have also been reviewed in various contexts and for different purposes (e.g., Borgman and Furner 2002; Bornmann and Daniel 2008; Tabatabaei 2013).

In order to assign different weights to citations of different importance automatically, with the goal of improving citation analysis and information retrieval results, studies have explored how textual properties, including citation frequency and citation location in the citing papers, may be used to automatically differentiate citations of different importance to citing papers. Since it has been consistently found that in-text frequency of a cited reference indicates its importance, frequency-based approaches assign a weight of N (or a function of N such as N^2) to a citation that appears N times in a citing paper (Bu et al. 2018; Zhu et al. 2015). Although some studies (e.g. Hanney et al. 2005) found no significant difference in terms of citation location for citation importance, many studies found that citations located in methodology, results, discussion, or conclusion sections may play a more significant or meaningful role than those located in introductory sections (Bertram 1972; Bonzi 1982; Cano 1989; Tang and Safer 2008; Voos and Dagaev 1976; Zhao et al. 2017). Location-based approaches would therefore assign more weight to citations in these sections. McCain and Turner (1989) experimented with weighing citations by a combination of their in-text frequency, location, and self-citation in an attempt to construct a "utility index" for citations. Considering both frequency and location in citation weighting can be more effective (Ding et al. 2013; Herlach 1978). For example, Herlach (1978) noted that a paper that has been cited in the Introduction or Literature Review, and subsequently mentioned in the Methodology or Discussion sections, will likely have made a more significant contribution to the citing article than one which has been mentioned only once in the entire article. Tang and Safer (2008) also emphasized other factors that may affect the impact of citation frequency on citation significance such as the "pond effect" (p. 262).

If the signal to be detected in citation analysis is the direct and substantial flow of knowledge from cited to citing papers, perfunctory citations can be considered a source of noise. This noise is quite strong as a high incidence of perfunctory citations has been repeatedly observed in previous studies. For example, Moravcsik and Murugesan (1975) noted that 40% references were perfunctory; Teufel et al. (2006) found that only a fifth of references are essential for the citing papers; Zhao et al. (2017) found that 65% of in-text citations were perfunctory or reviewed.

There are two approaches to dealing with noise: filter out the noise, or amplify the signal. The ultimately best approach is likely some combination of the two. All the frequency-based citation-weighting schemes found in the literature belong to the signal amplification approach. Compared to this approach, the noise filtration approach, which was introduced

by Zhao and Strotmann (2016), attempts to make the fundamental qualitative distinction between references that represent real use by, core impact on, or organic connection with, the citing paper (which it aims to retain for analysis) and those that are merely mentioned in passing as related work or background information (which it aims to remove). By only counting core connections in knowledge networks, this approach might help research evaluation become more sensitive to the essential impact of research. It could also better capture “aboutness” of documents, the essence of subject indexing in knowledge representation and retrieval. Knowledge representation and retrieval systems that make use of citation links could therefore benefit from improved precision in computer-aided subject indexing and in their “more like this” features (Zhao and Strotmann 2015b, 2016). In addition, the signal amplification required to counter the strong noise created by nonessential citations (65–80%) tends to be so strong (N^2 is the minimal power of N required) that it can cause serious distortions (Teufel et al. 2006; Zhao and Strotmann 2016). Filtering out most of this noise before applying the necessary signal amplification can avoid this technical problem.

The key, and difficult, question is how to identify and filter out perfunctory or nonessential citations. Zhao and Strotmann (2015b, 2016) proposed a simple method for this: re-citation analysis, which focuses on re-citations, i.e., references that appear more than once in the text of a citing paper, by filtering out uni-citations, i.e., references that appear only once in the text of a citing paper. The basic assumption of re-citation analysis is that papers are likely to be cited multiple times in a publication that relies heavily on them, while merely perfunctory citations should appear only once in a citing paper. Zhao et al. (2017) tested this assumption in the library and information science field and found that quite a large percentage (about 30%) of multi-citations play purely a nonessential role, while about 30% of uni-citations play an essential role in the citing paper. This means that 30% multi-citations would be over-weighted (false-positives) by frequency-weighted citation counting and 30% uni-citations would be disregarded unfairly (false-negatives) by re-citation analysis. It was found that removing citations by location could be more effective for filtering out nonessential citations than removing uni-citations, and that removing all citation occurrences in the Background and Literature Review sections, and all uni-citations in the Introduction section, appears to provide a good balance between filtration and error rates (Zhao et al. 2017).

The present study therefore examines what differences location-filtered citation counting might make in citation analysis for research evaluation. Specifically, the present study addresses the following research questions.

- (1) Does location-filtered citation counting make a significant difference in author rankings by citations compared to traditional citation counting?
- (2) What are some of the major differences?
- (3) Does location-filtered citation counting make essential impact stand out more?
- (4) What types of bibliometrics-related studies have been conducted in the biomedical fields?

Methodology

Our dataset for this study comprises the full text of all articles on bibliometric studies, especially citation analysis studies, available as full-text in PubMed Central (PMC). We chose a research area that we are knowledgeable about so that we are in a good position to

make sense of the results. PMC was chosen for its quality indexing, full text availability, and because it primarily covers journals in the biomedical research fields.

We conducted a search in PMC in April 2019 using the search string below and had 6088 hits.

```
"citation analysis" OR bibliometric
```

In order to focus on applications of these methodologies in the biomedical research fields, we removed articles published in the journal *Scientometrics* which publishes generic bibliometrics-focused studies. Considering how few articles in the search results were published in the journal *Scientometrics* (a total of 68) and the fact that PMC primarily covers journals in the biomedical research fields, it is safe to say that only a small percentage of articles in the search results might not be targeted to scientists in biomedical fields. The dataset collected this way should therefore suffice for the purpose of this study, given the acceptable 20% error rate in data that most if not all bibliometric studies work with (Strotmann and Zhao 2015).

We downloaded the full XML records of all articles that have full-text available and that have at least one cited reference as our dataset. This dataset has 3681 citing articles, which contained a total of 176,065 reference list entries, and a total of 265,444 in-text citation occurrences for these entries.

For each full text in this dataset, we counted the in-text citations to the first author of each cited reference in the following ways. The total citation count for an author (operationalized here as surname plus initials of first author listed in the PMC XML file for a reference) is then calculated as the sum over all distinct reference list entries with this author in all full-text articles in the dataset. Essentially, we used the paper-based (instead of author-based) counting approach discussed in Zhao and Strotmann (2016).

1. W1—this is traditional citation counting, which adds 1 to an author's citation count whenever a paper with this author listed as first author is cited, regardless of how many times this paper is cited in the text.
2. Wn—this method adds N to an author's citation count when a paper with this author listed as first author is cited N times in a citing text.
3. EssW1—Remove introductory and background sections (i.e., introduction, literature review, related studies, background) and then count W1.
4. EssWn—Remove introductory and background sections (i.e., introduction, literature review, related studies, background) and then count Wn.

The structure of scientific articles reporting original research results has been, to a large degree, standardized over the years to include "introduction," "methods," "results," "discussion," and "conclusion" sections (Doumont 2010; Thelwall 2019b). This structure reflects the progression of most research projects, facilitates more effective and efficient use of research articles, and has been recommended by many style manuals and required by most scientific journals (Doumont 2010; McCain and Turner 1989).

Based on previous studies, we attempted to filter nonessential citation by removing introductory and background information including the introduction, literature review, related studies, and backgrounds sections. The Introduction section has been found to be somewhat different from the other sections in that the percentage of nonessential citations was in the eighties there instead of nineties in the other sections. Ideally, only uni-citations in the introduction section and all citations in the other sections should be removed as this was found to provide a good balance between filtration and error rates (Zhao et al. 2017). We removed all citations

instead of just uni-citations from the Introduction section in this first attempt to test the differences that location-filtered citation counting makes in research evaluation, with the assumption that doing so would allow us to study one effect at a time—namely, in the present paper, that of filtering citations by the section type they occur in.

Author names were ranked by each of these above-listed counts in the usual way, using the average rank to number tied authors, i.e., all names with the same citation count are assigned the average of their ranks. In total, we computed four different rankings of over 82 thousands of first-author names that are cited in our dataset.

To examine how the various author rankings are different from each other, we first calculated the Spearman correlations of author rankings by these four methods for the 500 most highly cited authors selected by average rank over the four counting methods (Al Jaber and Elayyan 2018). We then examined rank changes of individual authors and the topics of their highly cited papers for the top 100 authors. For this detailed examination of rank changes, we manually removed author names that are likely ambiguous including all Chinese names.

We did not perform automatic author name disambiguation in any of the four counting methods we compared, and we only counted the first author of each cited reference. Performing automatic disambiguation and counting all authors might well change specific ranks of individual authors, but we cannot think of any reason why that would be able to drastically change the rank *difference* of the same author between two rankings. In particular, the very large or very small rank differences that we relied on in our analysis would remain large, or small, respectively, in practically all cases were disambiguation and all author counts used to determine the same four rankings instead.

Results and discussions

Correlation

Table 1 presents the Spearman correlations of rankings of top 500 cited author names. All correlations are significant at the 0.01 level (2-tailed).

It is interesting and somewhat surprising to see that the ranking by EssW1 has a fairly high correlation with the ranking by W1 (0.69), the traditional counting method, considering that a large percentage of citations were discarded when counting EssW1 (Bertin et al. 2016; Thelwall 2019a; Zhao et al. 2017).

The ranking by EssW1 has a similar correlation to the one by Wn with the traditional ranking (0.69 vs. 0.70). Since direct frequency-weighting (Wn) has been found to be insufficient to predict important citations compared to squared frequency-weighting (Zhu et al. 2015), it appears that simply removing article sections that contain mostly nonessential citations without applying a weighting scheme (EssW1) is probably just as insufficient.

However, the combination of the two, which filters out a large source of nonessential citations first and then weighs citations by their in-text frequency, makes a large difference

Table 1 Spearman correlations between rankings of top 500 authors

	W1	EssW1	Wn
EssW1	0.69		
Wn	0.70	0.52	
EssWn	0.40	0.64	0.77

in author ranking compared to traditional citation counting, as shown by the 0.4 correlation between W1 and EssWn. This indicator therefore deserves further investigation on whether it could help improve citation analysis results.

This result, i.e., filtering out likely nonessential in-text citations and then weighing the remaining ones by in-text frequency having a large difference from (i.e., low Spearman's rank correlation to) traditional counting, is similar to the one in Zhao and Strotmann (2015b, 2016) that proposed and tested filtering uni-citations. However, the present paper's method for identifying likely candidates for nonessential citations for filtering is likely to have a much higher accuracy rate than the one used in the earlier papers that filters out uni-citations (Zhao et al. 2017).

More and more studies have realized that citation impact is multidimensional and requires different indicators to measure its different dimensions (Bu et al. 2018; Funk and Owen-Smith 2017; Wu et al. 2019). Traditional citation counting appears to favor broad and shallow impact as indicated by findings from previous studies that in-text frequency of a cited reference indicates its importance to the citing paper on the one hand and uni-citations dominate and tend to be highly cited references on the other. EssWn appears to be able to emphasize deep and narrow impact, with its two ingredients of frequency-weighting and location-filtering leaning towards the deep aspect and the narrow (e.g., methodological contributions) aspect, respectively.

Author rank variability

Table 2 lists the top 100 author names selected by the average of their ranks assigned by all four counting methods, along with their ranks and the differences in ranks between traditional citation counting (W1) and weighted counting of citations (EssW1, EssWn, Wn). However, we will focus on comparing EssWn with W1, the most dissimilar, for this initial study, leaving the comparison of other methods to future studies. Author names in Table 2 are ranked by W1.

The variability of author ranks by these different counting methods is clearly visible. A general pattern seen from Table 2 for EssWn compared to W1 is that (a) ranks for highly cited authors are relatively stable, (b) rank gains occur mostly in the bottom 40% of the list, and (c) rank drops occur mostly in between.

- Authors with stable ranks

The most highly cited authors with stable ranks include both bibliometricians (e.g., Garfield, Bornmann, Leydesdorff, Van Eck, Falagas) and biomedical researchers (e.g., Moher, Sweileh, Zyoud, Huh), as well as authors with signal works that influenced bibliometrics highly (i.e., Hirsch).

A common feature of these bibliometricians is that they introduced, tested, and promoted methods, indicators, or tools for research evaluations and research network analysis—Hirsch's h-index, Van Eck's VOSviewer—a visualization tool for studying co-authorship networks, word co-occurrence networks and citation networks, Leydesdorff's work on the Triple Helix of university–industry–government relations, Glänzel's work on co-authorship analysis, Waltman's work on the Leiden ranking methods including the crown indicator, and Falagas' comparison of major citation databases. This feature is not all that surprising, considering that methodology sections are one of the sections that were specifically kept in calculating EssWn.

Table 2 Ranks and rank differences of top 100 authors

AuthorName	W1	EssW1	Wn	EssWn	W1-EssWn	W1-EssW1	W1-Wn
Garfield,E	1	2	1	3	-2	-1	0
Bornmann,L	2	6	2	5.5	-3.5	-4	0
Moher,D	3	1	3	1	2	2	0
Sweileh,WM	4	4	6	5.5	-1.5	0	-2
Hirsch,JE	5	7	4	7	-2	-2	1
Zyoud,SH	6	3	5	2	4	3	1
Falagas,ME	7	5	7	4	3	2	0
Huh,S	8.5	9	15	9	-0.5	-0.5	-6.5
Van Eck,NJ	11	8	12	8	3	3	-1
Leydesdorff,L	12	15	13.5	23	-11	-3	-1.5
Seglen,PO	13	13	13.5	24	-11	0	-0.5
Moed,HF	14	23.5	18	35	-21	-9.5	-4
Waltman,L	15	14	17	17.5	-2.5	1	-2
Masic,I	16.5	60	8	85.5	-69	-43.5	8.5
Groneberg-Kloft,B	18	10	21	12.5	5.5	8	-3
Egghe,L	19.5	25	23.5	33	-13.5	-5.5	-4
Glänzel,W	19.5	19.5	23.5	25.5	-6	0	-4
Eysenbach,G	22	29.5	19	25.5	-3.5	-7.5	3
Smith,R	23.5	83	33	149.5	-126	-59.5	-9.5
Larivière,V	23.5	27.5	20	29	-5.5	-4	3.5
López-Muñoz,F	25	36.5	29.5	60	-35	-11.5	-4.5
Merton,RK	26.5	32.5	35	62	-35.5	-6	-8.5
Boyack,KW	28	32.5	32	56	-28	-4.5	-4
Van Noorden,R	29	96	47	130	-101	-67	-18
Lewison,G	30.5	32.5	40.5	44	-13.5	-2	-10
Cramer,H	32	76	29.5	94.5	-62.5	-44	2.5
Kulkarni,AV	33.5	19.5	51	27.5	6	14	-17.5
Thelwall,M	36	66	40.5	99	-63	-30	-4.5
Van Raan,AFJ	36	43.5	49.5	71	-35	-7.5	-13.5
Greenhalgh,T	36	21	25	14	22	15	11
Hopewell,S	38	23.5	27	20	18	14.5	11
Altman,DG	39	27.5	36.5	37.5	1.5	11.5	2.5
Radicchi,F	40.5	60	46	94.5	-54	-19.5	-5.5
Fanelli,D	43	76	39	99	-56	-33	4
Schulz,KF	44	26	48	35	9	18	-4
Chalmers,I	45	83	65	139	-94	-38	-20
Bollen,J	46.5	122	62.5	157.5	-111	-75.5	-16
Yeung,AWK	49	38	87	94.5	-45.5	11	-38
Opthof,T	50.5	89	54.5	130	-79.5	-38.5	-4
Haustein,S	50.5	76	44.5	85.5	-35	-25.5	6
Small,H	53.5	114	95	250.5	-197	-60.5	-41.5
Buxton,M	53.5	56.5	22	22	31.5	-3	31.5
Costas,R	56	53.5	57	73.5	-17.5	2.5	-1
Cohen,J	58.5	32.5	62.5	45	13.5	26	-4

Table 2 (continued)

AuthorName	W1	EssW1	Wn	EssWn	W1-EssWn	W1-EssW1	W1-Wn
Börner,K	61.5	114	94	250.5	-189	-52.5	-32.5
Hicks,D	61.5	89	57	89.5	-28	-27.5	4.5
Adams,J	61.5	41	57	56	5.5	20.5	4.5
Uthman,OA	64.5	53.5	52	47.5	17	11	12.5
Barabási,AL	66	60	54.5	85.5	-19.5	6	11.5
Dickersin,K	67.5	46.5	101.5	94.5	-27	21	-34
Abramo,G	71.5	96	106.5	179.5	-108	-24.5	-35
Porter,AL	75.5	107	97.5	192.5	-117	-31.5	-22
Björk,BC	75.5	66	87	130	-54.5	9.5	-11.5
Katz,JS	75.5	141.5	82.5	130	-54.5	-66	-7
Callaham,M	78.5	36.5	87	52	26.5	42	-8.5
Wuchty,S	80.5	101	109	179.5	-99	-20.5	-28.5
Petersen,AM	80.5	114	65	81.5	-1	-33.5	15.5
Von Elm,E	87	52	106.5	94.5	-7.5	35	-19.5
Tadmouri,GO	87	56.5	109	85.5	1.5	30.5	-22
Bhandari,M	87	56.5	97.5	70	17	30.5	-10.5
Milat,AJ	87	83	49.5	43	44	4	37.5
Clauset,A	90	66	71.5	73.5	16.5	24	18.5
Watts,DJ	92	56.5	80	52	40	35.5	12
Grant,J	95	107	101.5	114	-19	-12	-6.5
Wooding,S	95	101	82.5	68	27	-6	12.5
Wagner,CS	98.5	76	112	107	-8.5	22.5	-13.5
Ramos,JM	98.5	66	77.5	60	38.5	32.5	21
Piwovar,HA	105.5	76	93	78	27.5	29.5	12.5
Fortunato,S	110	141.5	129	250.5	-140.5	-31.5	-19
Gastner,MT	110	41	109	52	58	69	1
Cooke,SJ	117	319	68	419.5	-302.5	-202	49
Horton,R	117	96	136	139	-22	21	-19
Benamer,HT	123.5	96	91	56	67.5	27.5	32.5
Michie,S	129.5	89	119.5	73.5	56	40.5	10
Arksey,H	129.5	50	61	31	98.5	79.5	68.5
Feixas,G	135	83	80	114	21	52	55
Graham,ID	135	132	87	114	21	3	48
Lavis,JN	139.5	163	97.5	192.5	-53	-23.5	42
Campbell,MK	139.5	101	145.5	166.5	-27	38.5	-6
Graffigna,G	146	96	71.5	64.5	81.5	50	74.5
Lefaiivre,KA	146	76	125.5	64.5	81.5	70	20.5
Kissin,I	153.5	122	156.5	192.5	-39	31.5	-3
Albert,R	153.5	107	174.5	192.5	-39	46.5	-21
Wasserman,S	153.5	107	165.5	149.5	4	46.5	-12
Banzi,R	153.5	107	73.5	68	85.5	46.5	80
Lee,J	160.5	132	129	114	46.5	28.5	31.5
Simonton,DK	166.5	205	65	419.5	-253	-38.5	101.5
Wager,E	166.5	163	87	107	59.5	3.5	79.5

Table 2 (continued)

AuthorName	W1	EssW1	Wn	EssWn	W1-EssWn	W1-EssW1	W1-Wn
Maziak,W	188.5	141.5	174.5	271	−82.5	47	14
Levac,D	188.5	122	174.5	121	67.5	66.5	14
Kuruvilla,S	199.5	150.5	59.5	27.5	172	49	140
Ponce,FA	213.5	183	174.5	229.5	−16	30.5	39
Bennett,S	213.5	122	148	89.5	124	91.5	65.5
Innes,KE	243	319	68	209.5	33.5	−76	175
Donner,A	243	141.5	199	209.5	33.5	101.5	44
Grimshaw,JM	243	205	161.5	157.5	85.5	38	81.5
Rogers,EM	243	163	199	119	124	80	44
Rudan,I	243	163	141.5	102.5	140.5	80	101.5
Clarke,M	243	319	141.5	102.5	140.5	−76	101.5
Kilkenny,C	278.5	141.5	223	121	157.5	137	55.5

The biomedical researchers with stable ranks (i.e., Sweileh, Zyoud, Huh) were cited for their actual bibliometric studies of biomedical fields as compared to those discussed below on the scholarly communication system in general and on problems in bibliometric indicators for biomedical research evaluation in particular. Moher was cited for the PRISMA statement (Preferred reporting items for systematic reviews and meta-analyses), CONSORT guidelines (for reporting parallel group randomised trials), and other guidelines of this sort. Moher being consistently ranked high indicates that a large part of our dataset is systematic reviews and meta-analyses, and that bibliometric methods have been used in these types of studies.

- Authors who have been ranked higher by EssWn than by W1 counting

Authors in the bottom 40% of the table are dominantly biomedical researchers and ranked much higher by EssWn. The largest gains in rank are represented by Kuruvilla, Kilkenny, Rudan, Clarke, Rogers, and Bennett at the very bottom of the table whose ranks are more than 120 places higher by EssWn than by W1. These authors stand out much more after authors, primarily bibliometricians, whose work was primarily cited in the introductory and background sections were pushed down the list.

Many of these authors provided theories or methodologies for doing various kinds of health research, which aligns very well with the emphasis of EssWn on the methodology, discussions, and conclusions sections of citing papers. For example, Kuruvilla and Banzi proposed conceptual frameworks for describing or assessing the impact of health research; Clarke promoted standardization of reporting outcomes for clinical trials and systematic reviews; Kilkenny was mostly cited for his work on the ARRIVE guidelines for reporting animal research; Arksey and Levac worked towards methodological frameworks for scoping studies. Their being ranked much higher by EssWn than by W1 indicates that systematic reviews and scoping studies were among the kinds of studies that have dominated bibliometrics-related studies in biomedical fields, just as was found from examining authors with stable ranks above.

Rogers was not a biomedicine researcher but was nevertheless pushed much higher (by 124 places) by citations to his theory on the diffusion of innovations after introductory and

background sections were removed. Watts was also pushed higher (by 40 places) by citations to his work on small world networks. It appears that their theories were actually used, instead of just being mentioned as background information, e.g., to explore and make sense of knowledge translation patterns in the biomedicine fields.

As would be expected theoretically, EssWn counting does appear to give advantage to authors who have done theoretical or methodological work and to give more weight to studies primarily concerning biomedical issues.

- Authors whose ranks dropped by EssWn counting

A striking pattern here is that all authors whose work was not focused on biomedical research have been ranked lower after introductory and background sections were removed and supposedly medical related contents were kept. These authors included almost all bibliometricians and researchers in the biomedical fields or other fields who were interested in scholarly communication in general and problems in research evaluation and science communication in biomedical fields in particular. Their drop in rank indicates that their works have been mostly cited as background information.

Among the biomedical researchers, the largest drops are represented by Smith, Van Noorden, Chalmers, Opthof, and Masic (although Van Noorden is a senior news editor for *Nature*). Masic and Van Noorden were highly cited for their ideas on problems in scholarly communication and publishing in biomedical fields, while Smith and Opthof appeared to have been mostly cited for their work published in medical journals in fields such as epidemiology and cardiology, on problems with the journal impact factor.

Most bibliometricians had small to medium drops in rank. The largest drops are represented by Small, Börner, Porter, Bollen, Abramo, and Thelwall (by 197, 189, 117, 111, 108 and 63 places respectively). Small and Börner are known for their work on science mapping using co-citation analysis and other methods. As discussed above, bibliometrics-related studies in the biomedical fields appear to be mostly systematic reviews, scoping studies, and meta-analysis type of studies. The large drops in Small and Börner's ranks from W1 to EssWn indicates that science mapping in general and co-citation analysis in particular have not been used much but were mostly mentioned as background information in these studies, which is unfortunate because co-citation networks and other citation-based network analysis methods (e.g., bibliographic coupling analysis) can be very informative of intellectual structures of research fields (e.g., White and McCain 1998; Zhao and Strotmann 2008a, b, 2011, 2014). Thelwall is a highly cited webometricians and Bollen has been cited mostly for his work on alternative impact measures (called Altmetrics) compared to citation impact metrics. Their work was not actually used in the studies of biomedical fields but was mostly mentioned as background information, indicating that bibliometrics-related studies in biomedical fields were mostly evaluative (as opposed to relational) studies based on citation measures (as opposed to altmetrics) of journal articles (as opposed to websites) even though the authors were aware of limitations imposed by this focus.

All three types of rank changes described above show that authors whose cited articles had a biomedicine focus rank higher after introductory and background sections were removed, whereas those with an emphasis on bibliometrics or scholarly communication rank lower. Considering that bibliometric studies in the biomedicine fields are mostly concerned with biomedicine, this general pattern makes good sense and indicates that EssWn weighs citations appropriately, i.e., assigning greater weight to essential citations than to perfunctory ones, and that citation analysis based on this indicator may indeed be able to focus on more profound influences and organic relationships.

Conclusions

It has been found that in-text frequency of a cited reference indicates its importance but a significant percentage of multi-citations are nonessential citations, and could be over-weighted by frequency-weighted citation counting. It is therefore important to filter out nonessential in-text citations before assigning weight in order to improve the accuracy and effectiveness of frequency-weighted citation analysis.

Previous studies proposed and tested filtering nonessential citations by their in-text frequency, assuming that uni-citations are mostly nonessential (Zhao and Strotmann 2015b, 2016), but found that its error rate might be too high (Zhao et al. 2017). The present study explores an alternative filtering method to see what a difference it might make in ranking authors by citations. Informed by findings from previous studies that citations located in methodology, results, discussion, or conclusion sections may play a more significant or meaningful role than those located in introductory and background sections, we removed introductory and background sections as a way to filter out nonessential citations, which was found to likely have a lower error rate and a higher filtration rate compared to removing uni-citations (Zhao et al. 2017). We analyzed the correlations between author rankings by traditional citation counting and those by in-text frequency-weighted citation counting before and after the filtration, and examined rank changes of individual authors between the traditional citation counting and location filtered counting weighted by in-text frequency.

We found that removing introductory and background sections alone does not make much of a difference in author rankings, but it makes a large difference when combined with frequency-weighted counting. This combination appears to make essential citations stand out, as shown by its generally ranking biomedicine-focused authors higher and bibliometrics-focused ones lower. Citation impact is multidimensional and requires different indicators to measure its different dimensions (Bu et al. 2018; Funk and Owen-Smith 2017; Wu et al. 2019; Zhao and Strotmann 2016). Traditional citation counting appears to (over-)emphasize broad but shallow impact whereas location-filtering citation counting weighted by in-text frequency appears to be able to focus on deep and narrow impact.

The present study also finds that weighing and filtering appear to have different effects on citation counting as indicated by a medium correlation between rankings by each separately. This difference warrants future detailed studies to identify the separate factors involved. Theoretically, they appear to each lean towards a different dimension of the deep and narrow impact that their combination appears to measure.

The observation that authors who represent guidelines for doing systematic review, reporting meta-analysis results, or conducting scoping studies are ranked much higher by EssWn counting indicates that many articles retrieved from PMC on bibliometrics in general and on citation analysis in particular belong to these types of studies. While bibliometric methods and tools have clearly been used in these studies, this use, however, didn't appear to have included citation-based knowledge network analysis methods such as co-citation analysis or bibliographic coupling analysis. Such methods have been shown to effectively reveal intellectual structures of research fields, and would be expected to be very useful for systematic reviews and scoping studies. It should be an interesting future study to find out why they have not been applied as much in bibliometric studies of biomedical fields.

References

- Al Jaber, A.O., & Elayyan, H.O. (2018). *Toward quality assurance and excellence in higher education*. Gistrup: River Publishers. p. 284.
- Bertin, M., & Atanassova, I. (2018). InTeReC: In-text reference corpus for applying natural language processing to bibliometrics. Retrieved April 18, 2019 from <http://ceur-ws.org/Vol-2080/paper6.pdf>.
- Bertin, M., Atanassova, I., Gingras, Y., & Larivière, V. (2016). The invariant distribution of references in scientific articles. *Journal of the Association for Information Science and Technology*, 67(1), 164–177.
- Bertram, S. (1972). Citations counts. In A. Pitelick (Ed.), *Proceedings of the fourth annual meeting of the American society for information science, Western Canada chapter* (pp. 61–67). Vancouver: University of British Columbia.
- Bonzi, S. (1982). Characteristics of a literature as predictors of relatedness between cited and citing works. *Journal of the American Society for Information Science*, 33, 208–216.
- Borgman, C. L., & Furner, J. (2002). Scholarly communication and bibliometrics. *Annual Review of Information Science and Technology*, 36, 3–72.
- Bornmann, L., & Daniel, H.-D. (2008). What co-citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, 64, 45–80.
- Boyack, K. W., Small, H., & Klavans, R. (2013). Improving the accuracy of co-citation clustering using full text. *Journal of the American Society for Information Science and Technology*, 64(9), 1759–1767.
- Boyack, K. W., Van Eck, N. J., Colavizac, G., & Waltman, L. (2018). Characterizing in-text citations in scientific articles: A large-scale analysis. *Journal of Informetrics.*, 12(1), 59–73.
- Brooks, T. A. (1985). Private acts and public objects: An investigation of citer motivations. *Journal of the American Society for Information Science*, 36(4), 223–229.
- Brooks, T. A. (1986). Evidence of complex citer motivations. *Journal of the American Society for Information Science*, 37(1), 34–36.
- Bu, Y., Waltman, L., and Huang, Y. (2018). A multidimensional perspective on the citation impact of scientific publications. Retrieved April 2018 from.
- Cano, V. (1989). Citation behavior—classification, utility, and location. *Journal of the American Society for Information Science*, 40, 284–290.
- Case, D. O., & Higgins, G. M. (2000). How can we investigate citation behavior? A study of reasons for citing literature in communication. *Journal of the American Society for Information Science*, 51(7), 635–645.
- Chubin, D. E., & Moitra, S. D. (1975). Content analysis of references: Adjunct or alternative to citation counting? *Social Studies of Science*, 5(4), 423–441.
- Ding, Y., & Cronin, B. (2011). Popular and/or prestigious? Measures of scholarly esteem. *Information Processing and Management*, 47, 80–96.
- Ding, Y., Liu, X., Guo, C., & Cronin, B. (2013). The distribution of references across texts: Some implications for citation analysis. *Journal of Informetrics*, 7(3), 583–592.
- Doumont, J. (Ed.). (2010). *English communication for scientists*. Cambridge: NPG Education. Retrieved September 22, 2016, from <http://www.nature.com/scitable/ebooks/english-communication-for-scientists-14053993>.
- Funk, R. J., & Owen-Smith, J. A. (2017). Dynamic network measure of technological change. *Management Science*, 63, 791–817.
- Garfield, E. (1979). *Citation indexing—its theory and application in science, technology, and humanities*. New York: Wiley.
- Hanney, S., Frame, I., Grant, J., Buxton, M., Young, T., & Lewison, G. (2005). Using categorizations of citations when assessing the outcomes of health research. *Scientometrics*, 65, 357–379.
- Herlach, G. (1978). Can retrieval of information from citation indexes be simplified? Multiple mention of a reference as a characteristic of the link between cited and citing article. *Journal of the American Society for Information Science*, 29(6), 308–310.
- Hou, W., Li, M., & Niu, D. (2011). Counting citations in texts rather than reference lists to improve the accuracy of assessing scientific contribution. *BioEssays*, 33, 724–727.
- Hsiao, T.-M., & Chen, K.-H. (2018). How authors cite references? A study of characteristics of in-text citations. *ASIST Proceedings*, 55(1), 179–187.
- Hu, Z., Lin, G., Sun, T., & Hou, H. (2017). Understanding multiply mentioned references. *Journal of Informetrics*, 11(4), 948–958.
- Jeong, Y. K., Song, M., & Ding, Y. (2014). Content-based author co-citation analysis. *Journal of Informetrics*, 8(1), 197–211.

- Liu, M. (1993). The complexities of citation practice: A review of citation studies. *Journal of Documentation*, 49, 370–408.
- McCain, K. W., & Turner, K. (1989). Citation context analysis and aging patterns of journal articles in Molecular-Genetics. *Scientometrics*, 17, 127–163.
- Moravcsik, M. J., & Murugesan, P. (1975). Some results on the function and quality of citations. *Social Studies of Science*, 5(1), 86–92.
- Narin, F. (1976). *Evaluative bibliometrics: The use of publication and citation analysis in the evaluation of scientific activity*. Washington, DC: Computer Horizons.
- Otto, W., Ghavimi, B., Mayr, P., Piryani, R. & Singh, V.K. (2019). Highly cited references in PLOS ONE and their in-text usage over time. Retrieved April 18, 2019 from ArXiv.org.
- Pak, C., Yu, G., & Wang, W. (2018). A study on the citation situation within the citing paper: Citation distribution of references according to mention frequency. *Scientometrics*, 114(3), 905–918. <https://doi.org/10.1007/s11192-017-2627-0>.
- Shadish, W. R., Tolliver, D., Gray, M., & Gupta, S. K. S. (1995). Author judgements about works they cite: Three studies from psychology journals. *Social Studies of Science*, 25(3), 477–498.
- Small, H. (1982). Citation context analysis. In B. J. Dervin & M. J. Voigt (Eds.), *Progress in communication sciences*, 3 (pp. 287–310). Norwood: Ablex.
- Strotmann, A., & Zhao, D. (2015). An 80/20 data quality law for professional scientometrics? In *Proceedings of the 15th International Society for Scientometrics and Informetrics Conference*, June 30–July 3, 2015, Istanbul, Turkey.
- Tabatabaei, N. (2013). *Contribution of information science to other disciplines as reflected in citation contexts of highly cited JASIST papers*. Montreal. (McGill University Ph.D. dissertation).
- Tang, R., & Safer, M. A. (2008). Author-rated importance of cited references in biology and psychology publications. *Journal of Documentation*, 64, 246–272.
- Teufel, S., Siddharthan, A., & Tidhar, D. (2006). Automatic classification of citation function. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing* (pp. 103–110). Stroudsburg.
- Thelwall, M. (2019a). Should citations be counted separately from each originating section? *Journal of Informetrics*, 13, 658–678.
- Thelwall, M. (2019b). The rhetorical structure of science? A multidisciplinary analysis of article headings. *Journal of Informetrics*, 13(2), 555–563.
- Vinkler, P. (1987). A quasi-quantitative citation model. *Scientometrics*, 12(1), 47–72.
- Voos, H., & Dagaev, K. S. (1976). Are all citations equal? Or Did we op. cit. your idem? *Journal of Academic Librarianship*, 1, 20–21.
- White, H. D. (1990). Author co-citation analysis: Overview and defense. In C. L. Borgman (Ed.), *Scholarly communication and bibliometrics* (Vol. 84, p. 106). Newbury Park: Sage.
- White, H. D., & McCain, K. W. (1998). Visualizing a discipline: An author co-citation analysis of information science, 1972–1995. *Journal of the American Society for Information Science*, 49, 327–355.
- White, H. D., & Wang, P. L. (1997). A qualitative study of citing behavior: contributions, criteria, and meta-level documentation concerns. *Library Quarterly*, 67, 122–154.
- Wu, L. F., Wang, D. S., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. *Nature*, 566, 378–382.
- Zhao, D., Cappello, A., & Johnston, L. (2017). Functions of uni-and multi-citations: Implications for weighted citation analysis. *Journal of Data and Information Science*, 2(1), 51–69.
- Zhao, D., & Strotmann, A. (2008a). Information Science during the first decade of the Web: An enriched author co-citation analysis. *Journal of the American Society for Information Science and Technology*, 59(6), 916–937.
- Zhao, D., & Strotmann, A. (2008b). Evolution of research activities and intellectual influences in Information Science 1996–2005: Introducing author bibliographic coupling analysis. *Journal of the American Society for Information Science and Technology*, 59(13), 2070–2086.
- Zhao, D., & Strotmann, A. (2011). Intellectual structure of Stem Cell research: A comprehensive author co-citation analysis of a highly collaborative and multidisciplinary field. *Scientometrics*, 87(1), 115–131.
- Zhao, D., & Strotmann, A. (2014). The knowledge base and research front of Information science 2006–2010: An author co-citation and bibliographic coupling analysis. *Journal of the Association for Information Science and Technology*, 65(5), 996–1006.
- Zhao, D., & Strotmann, A. (2015a). Analysis and visualization of citation networks. In Gary Marchionini (Ed.), *Synthesis lectures on information concepts, retrieval, and services*. San Rafael: Morgan & Claypool Publishers. <https://doi.org/10.2200/s00624ed1v01y201501icr039>.

- Zhao, D. & Strotmann, A. (2015b). Re-citation analysis: Promising for research evaluation, knowledge network analysis, knowledge representation and information retrieval? In *Proceedings of the 15th International Society for Scientometrics and Informetrics Conference*, June 30–July 3, 2015, Istanbul, Turkey.
- Zhao, D., & Strotmann, A. (2016). Dimensions and uncertainties of author citation rankings: Lessons learned from frequency-weighted in-text citation counting. *Journal of the Association for Information Science and Technology*, 67(3), 628–671.
- Zhu, X., Turney, P., Lemire, D., & Vellino, A. (2015). Measuring academic influence: Not all citations are equal. *Journal of the Association for Information Science and Technology*, 66(2), 408–427.