A Comparison of Citer and Citation-Based Measure Outcomes for Multiple Disciplines

Isola Ajiferuke
Faculty of Information and Media Studies, University of Western Ontario, London, ON, Canada N6A 5B7. E-mail: iajiferu@uwo.ca

Kun Lu and Dietmar Wolfram
School of Information Studies, University of Wisconsin-Milwaukee, P.O. Box 413, Milwaukee, WI, USA 53201. E-mail: {kunlu, dwolfram}@uwm.edu

Author research impact was examined based on citer analysis (the number of citers as opposed to the number of citations) for 90 highly cited authors grouped into three broad subject areas. Citer-based outcome measures were also compared with more traditional citation-based measures for levels of association. The authors found that there are significant differences in citer-based outcomes among the three broad subject areas examined and that there is a high degree of correlation between citer and citation-based measures for all measures compared, except for two outcomes calculated for the social sciences. Citer-based measures do produce slightly different rankings of authors based on citer counts when compared to more traditional citation counts. Examples are provided. Citation measures may not adequately address the influence, or reach, of an author because citations usually do not address the origin of the citation beyond self-citations.

Introduction

For decades, citations have represented a fundamental unit of measure for assessing the influence of authors and their scholarly works. The tradition of citation has an even longer history. Garfield (1955) and Shapiro (1992) point to the use of references and citations before the 20th century. It was, however, with the development of citation indexes for the sciences, thanks to Eugene Garfield's efforts, that the large scale assessment of researchers and publication sources could become practical.

Clearly, citations represent an important form of recognition. The act of citing acknowledges the contributions made in previous works that the citing work builds on, and serves as an affirmation (or refutation) of the cited work. Citation counts continue to be controversial because of the varied reasons for citer motivation (Garfield, 1965; Smith, 1981; Brooks, 1985; MacRoberts & MacRoberts, 1989; White & Wang, 1997; Bornmann & Daniel, 2008), how citations are counted (Cronin & Overfelt, 1993), the varying degrees of influence of the cited work, the influence of uncited works that go unrecognized (MacRoberts & MacRoberts, 2010), and because citations may not address the quality of a research contribution (Lindsey, 1989; Phelan, 1989; Moed, 2005). MacRoberts and MacRoberts also noted that citing behavior can be biased, and self-citation (i.e., when authors cite themselves) can be common (Snyder & Bonzi, 1998). Ultimately, the act of citing is a subjective decision made by the citing author(s), and in the absence of additional evidence citations must be treated uniformly. Despite these issues a citation remains an important unit of measure because it defines relationships among works and authors. High citation counts are thought to provide evidence of the influence or significance of a given work, an author associated with the work, or a higher unit of aggregation such as a research group.

The purpose of the present research is to investigate the largely unexplored idea of citer analysis across different disciplines and to compare measures developed for citer analysis by the investigators with those based on more traditional citation analysis. In an earlier study, Ajiferuke and Wolfram (2010) examined the feasibility of citer analysis by performing a comparison of prolific authors in library and information science (LIS) from the United States and the United Kingdom. By focusing on the number of citers instead of citations, the investigators could more objectively determine the "reach" of an author's work. Citation-based research rarely considers the origin of the citation or the influence of recitations (the repeated citation of the same work by the same author). Whether a cited author receives 10 citations

Received March 3, 2010; revised May 7, 2010; accepted May 7, 2010

 • Published online 17 June 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/asi.21383

from one citing author or one citation each from 10 different citing authors is not considered in citation analysis—neither is the extent to which citation and citer-based counts differ. Our earlier research revealed that there were high correlations between citer and citation-based measures in most cases, but with some differences in author rankings. In the current study the authors examine several disciplines from the sciences and social sciences to determine whether the observed correlations for LIS are also found in the other disciplines. More specifically, the focus of the present research is to determine:

1. If there are significant differences in citer patterns across disciplines.
2. The extent to which citer and citation-based measures for highly cited authors in different social science and science disciplines are correlated.

Literature Review

Citation-based research is both longstanding and substantial. The use of a citation as the unit of analysis has been employed in studies that examine the impact of authors and publication sources, and has been applied in the development of measures such as the journal impact factor (Garfield & Sher, 1963) and the h-index for author research impact (Hirsch, 2005). Similarly, cocitation analysis, which explores associations among entities of interest based on citations, has been used to examine relationships between documents (Marshakova, 1973; Small, 1973), authors based on the visualization of cocitations (White & McCain, 1998), and the discovery of "undiscovered" knowledge based on the absence of cocitations (Swanson, 1986).

The majority of citation analysis studies have focused on the number of citations received but not the origin of those citations, except when considering citer motivation or self-citations. Between Garfield's (1965) initial assessment of citer motivation to Bornmann and Daniel's (2008) recent review of citing behavior and associated research, there have been many papers published outlining reasons for the citation of given works. Despite the recognition of the importance of the citer, few studies to date have relied on direct citer-based data as the focus of investigation. Dieks and Chang (1976) recognized the origin of citations by mathematically modeling citation processes that included the number of citers of papers. White (2001) examined the citing patterns for eight prominent information scientists through citation identities (i.e., whom the author of interest cites) and citation image makers (i.e., who cites the author of interest). His investigation revealed that the practice of reciting an author was common, with 49% or more of citations constituting recitations for six of the eight authors he examined. White also drew attention to the equivalency of citers and citations as types and tokens, respectively, and noted that most research to date has been conducted at the token level. Following White's work, Cronin and Shaw (2002) explored the citation identities for three information science scholars. They included the total number of citers in their report in addition to more traditional citation-based measures. Other studies have also examined the idea of citers as citation image makers for a specific author. These have included Eugene Garfield (White, 2000), Homi Jehangir Bhabha (Swarna, Kalyane, & Kumar, 2008), and Michael O. Rabin (Bar-Ilan, 2006). The citer data in these studies were reported as frequency distributions.

More recently, Katsaros, Akritidis, and Bazanis (2009) proposed an f index, which can be used to assign weights or values to citations. The authors also pointed to the need to acknowledge the origin of a citation. In particular, they introduced the concept of a coterminal citation, where an author has authored or coauthored multiple papers that cite the same paper. The authors discussed further how this relationship could be used in the assessment of the impact of a given work or author.

In our earlier research (Ajiferuke & Wolfram, 2010) we investigated the practicality of citer-based analysis on more than a few authors by performing a citer analysis of 51 highly cited researchers in LIS. The research was prompted by the question of whether the reach of an author's research is more accurately determined by the number of citations received (tokens), or the number of people (types) who have cited and been influenced by that author's work, that is, the number of (unique) citers. We noted the high correlations between citers and citations and their complementary nature. Although highly correlated in most cases, the use of citations versus citers could affect rankings of individuals due to potentially large differences between these values. The earlier study also revealed some of the challenges associated with citer-based measures, which have parallels in citation-based measures. Most notable is the influence of different levels of collaboration that may result in hyperauthorship (Cronin, 2001), where citing publications are authored by many people. Are all authors of a publication with 50 authors equally influenced by a cited work? Certainly this is unlikely to be the case. But there is no clear delimiter by which to decide that some authors of a publication are more likely to have been influenced than other authors because authors may be listed in an order that is based on different conventions (e.g., alphabetical, contributed effort, grouped by institutional affiliation, senior author listed last). Similarly, if a cited publication is authored by more than one person, should each author receive full or partial credit? We argue in favor of equal credit due to the abstract outcomes that result from partial credit for having influenced a citing author. A more detailed rationale for full credit can be found in Ajiferuke and Wolfram (2010).

Our earlier study represented an initial investigation examining only one discipline in which these issues were readily apparent. This research builds on the ideas of the initial work by including a larger number of disciplines and additional areas of focus.

Methodology

Nine social science and science fields included in Thomson/ISI's Web of Science (WoS) highlycited.com were identified for investigation (Economics/Business, Psychology/Psychiatry, Social Sciences–General, Engineering, Space Sciences, Mathematics, Chemistry, Clinical Medicine, Biosciences-Microbiology). Ten randomly selected researchers were identified from each field, totaling 90 authors. Data used for the study were limited to the period 1987–2007. All publication types were considered except book reviews, which were excluded. The authors recognize that the inclusion publication formats that are less likely to be cited, such as meeting abstracts, could affect measures based on average counts across publications. However, they will cast a broader net for identifying citations and citers. These more minor forms of publication typically represent a small percentage of the authors' publications.

For the first five fields, which contained manageable numbers of publications per author, data were tabulated using the Analyze Results feature of WoS for each publication receiving at least one citation. Data for each author's publications were stored and tabulated in MS Excel. The remaining four subject areas represented fields with very prolific authors and high levels of collaboration for which manual data analysis was impractical. Datasets for the authors of interest from these fields were obtained from Thomson Reuters and were analyzed using database management software and developed programs. As with other author-based citation studies, author disambiguation and scatter arising from homonyms and multiple spellings can be problematic. Some verification of authors of interest was performed at the publication level (e.g., through their affiliations) for the five fields processed manually, but citer identities were not verified. In this case the "Analyze Results" function of WoS was used. For the remaining four fields, publication data for specific authors along with links for cited and citing publications based on document identifiers were provided by Thomson Reuters. The document identifiers and automatic matching of author names were used to tabulate citing author and publication relationships of interest.

Citer and citation measure calculations were made at the individual publication level. The number of citations for each publication represented the tally of unique publications citing the publication of interest. Self-citations were removed. The definition of self-citation was extended to include citations originating from publications authored by one of the coauthors of the cited publication of interest, or coauthor self-citations, as outlined by Costas, van Leeuven, and Bordon (2010). For example, if publication P1, coauthored by authors A & B, is cited by publication P2, which is coauthored by authors B & C, this is considered a form of self-citation, even if author A is the author of interest. As a result, in calculating the number of citations, the citation made by publication P2 to P1 is not counted. However, publication P2 is handled a bit differently in calculating the citer count. Author B of publication P2 is a self-citer of P1 and is, therefore, not included in the citer count of author A. Author C, however, is counted as a valid citer for authors A or B, assuming that author C had not already been counted as a citer of publication P1 previously. This allows collaborators, who would otherwise be excluded as part of a self-citation, to be included in the citer count. Note that a given citer may be counted more than one time if the citing author is citing a different work by the cited author. This acknowledges that the cited author may have multiple contributions that have influenced citers; otherwise, a citer would only be counted once whether influenced by one work of the cited author or many such works.

The following measures were tabulated from the raw citation data (see the Appendix for the calculation of these measures for a sample author dataset):

> Adjusted Citation Count: number of citations received by an author minus self-citations.
> Adjusted Citations per Publication: total adjusted number of citations received by an author divided by the number of publications by the author.
> h-index: the classic definition of x publications with at least x citations (Hirsch, 2005) is used.
> Adjusted h-index: based on the adjusted citation count for the author.
> Citer Count: number of authors who have cited a publication by given author.
> Citers per Publication: number of citers divided by the number of publications by an author.
> Reciter Rate: the proportion of unique citers citing an article more than once.
> ch-index: an author's ch-index of value x corresponds to x publications with at least x citers (i.e., the citer equivalent of the h-index).

To permit statistically valid comparisons to be carried, data were grouped into three broad subject areas: Social Sciences (Economics/Business, Psychology/Psychiatry, Social Sciences–General), Mathematical/Engineering Sciences (Engineering, Space Sciences, Mathematics), and Biological/Medical Sciences (Chemistry, Clinical Medicine, Biosciences–Microbiology), with 30 authors per group.

The data provided by Thomson Reuters for Mathematics, Chemistry, Clinical Medicine, and Biosciences–Microbiology made it possible to investigate citer characteristics in more detail, including citer patterns over time. Because citation counts on their own do not reveal the origin of the citations, a citer analysis over time can show if a given work is continuing to attract new citers over time or if recitation is occurring.

Results

A summary of the descriptive outcomes for the 90 selected authors within each field appears in Table 1. The widely varying values across the different fields and within the three broad areas merits mention. Because Psychology/Psychiatry can overlap with clinical medicine, which exhibits the highest levels of citation and citer activity, it is not surprising that it would have higher values than the other social sciences. More detailed analysis for the three topic areas follows.

DOI: 10.1002/asi

TABLE 1. Summary of measures for each field.

Field	Total citations	Adjusted citation count	Citer count	Citations per pub.	Adjusted citations per pub.	Citers per pub.	h-index	Adjusted h-index	ch-index	Average reciter rate*	Average number of authors per publication
	Averages within fields										
Economics/Business	2045.5	1947.6	3132.2	51.4	49.0	80.1	20.7	20.0	23.8	0.1046	2.3
Psychology/Psychiatry	4538.1	4092.1	11197.7	43.7	39.8	104.0	31.7	30.2	45.9	0.1059	3.6
Social Sciences – General	1502.0	1369.1	3568.6	25.1	23.1	57.7	17.9	17.0	24.0	0.0873	3.1
Engineering	1383.6	1169.2	2694.4	24.2	20.2	45.4	18.4	16.1	22.9	0.1402	3.1
Mathematics	1872.2	1530.8	2991.0	32.1	27.5	58.0	20.1	17.9	22.8	0.1787	2.4
Space Sciences	6668.4	5698.4	15803.6	82.6	71.7	177.4	36.9	32.4	52.2	0.1844	9.0
Biosciences – Microbiology	8866.4	7578.9	29633.6	76.8	65.5	251.0	48.1	42.2	76.7	0.1588	4.7
Chemistry	9797.4	8264.6	22598.6	50.8	42.5	117.6	49.0	43.8	76.2	0.1779	3.6
Clinical Medicine	21818.8	19562.1	85168.5	102.3	92.3	396.0	72.9	67.5	127.7	0.1241	6.3

*The Average Reciter Rate is calculated by averaging the article reciter rate at the author level.

Number of Unique Citers

Descriptive statistics for the consolidated data for the three broad topic areas were analyzed to determine whether parametric or nonparametric statistics would be appropriate for the analysis of differences (Table 2). A Kruskal–Wallis test (Kvam & Vidakovic, 2007, Ch. 8) for differences between three or more groups was used to test for significant differences in the average citers across fields because of the skewness of the data. The outcome of the test is significant ($\chi^2 = 42.671$, $df = 2$, $p = 0.000$), with Biological/Medical Sciences having the highest median number of citers per researcher. Note that instances of hyperauthorship may unduly increase the number of citers recorded. In one case, more than 500 individuals were listed as the authors of a citing paper in Clinical Medicine. This, however, represented an anomalous and extreme case.

Intuitively, one would expect the number of citers to be higher than the number of citations because each author of a citing paper can contribute to the citer count instead of just contributing a single citation count for the paper. This was indeed the case for almost all researchers studied. However, there is the potential for the adjusted citation count to be higher than the number of citers. This likelihood is influenced by the average number of authors for citing articles, self-citations, and levels of recitation. This can be seen in the Social Sciences – General subject area, where researcher SS6 has a slightly higher adjusted citation count than citer count (Figure 1). For this researcher, the average number of authors per citing document was very low, with an average of 1.6 authors per citing article. Note, however, that in fields with higher levels of collaboration, like Clinical Medicine, which also attracts more citations, the differences between the adjusted number of citations and citers become quite large for highly cited authors (Figure 2).

Number of Citers per Publication

At least one of the three distributions is skewed, so a Kruskal–Wallis test was used to compare the distribution of values (Table 3). Again, there was a significant difference in the outcome ($\chi^2 = 31.256$, $df = 2$, $p = 0.000$), with Biological/Medical Sciences having the highest median number of citers per publication.

TABLE 2. Descriptive statistics for unique citers.

	Mean	Median	Standard deviation	Mean rank
Mathematical/Engineering Sciences	7163	5136	9944.110	33.27
Social Sciences	5966.167	3799	6313.431	32.30
Biological/Medical Sciences	45800.233	36358	39189.409	70.93

Reciter Rate

The proportion of reciters will influence the difference between citer and citation counts. Descriptive statistics appear in Table 4a. All distributions were fairly normal. A one-way analysis of variance (ANOVA) test to test for differences in mean values between three or more groups was conducted along with Ryan (i.e., REGWQ) multiple comparison tests to determine where any significant differences existed. The results in Table 5b indicate a significant difference, with the Social Sciences having significantly lower levels of recitation when compared to the other two groups of disciplines.

ch-Index

Descriptive statistics for the ch-index values for the three broad subject areas appear in Table 5a. The distributions again were fairly normal, so a one-way ANOVA was used (Table 5b). The Ryan multiple comparison tests indicate the Biological/Medical Sciences have much higher ch-index values than for the other two areas.

DOI: 10.1002/asi

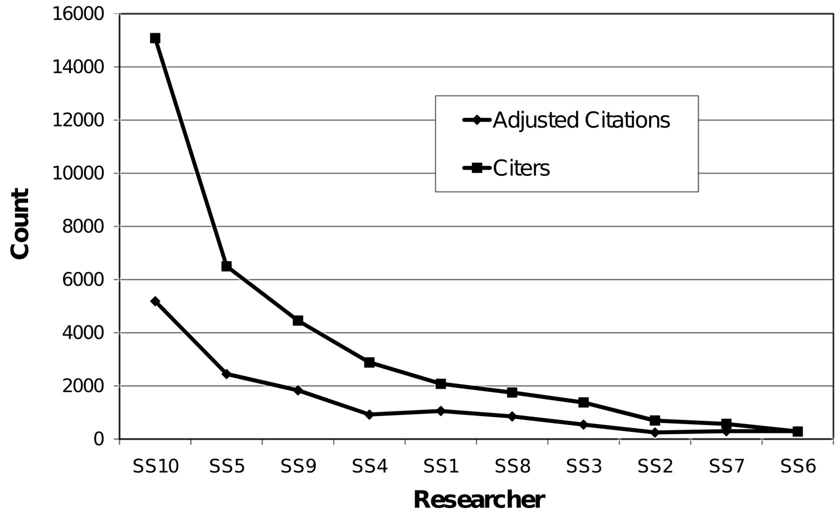

FIG. 1. Adjusted citation and citer count comparison for Social Sciences–General.

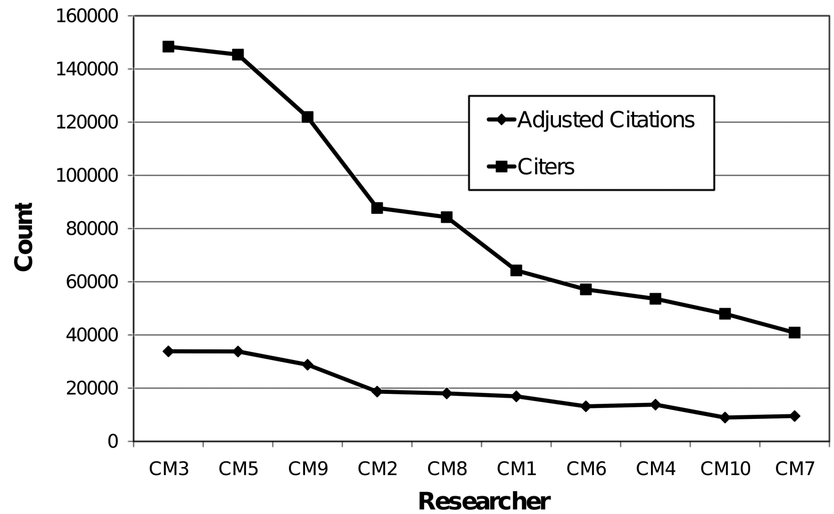

FIG. 2. Adjusted citation and citer count comparison for Clinical Medicine.

TABLE 3. Descriptive statistics for number of unique citers per publication.

	Mean	Median	Standard deviation	Mean rank
Mathematical/Engineering Sciences	93.6037	62.1979	99.541	33.50
Social Sciences	80.5749	78.7919	45.786	35.77
Biology/Medical Sciences	254.8508	192.747	182.54	67.23

TABLE 4a. Descriptive statistics for reciter rate.

	Min	Max	Mean	Median	St. dev.
Mathematical/Engineering Sciences	0.09	0.33	0.1678	0.16578	0.04658
Social Sciences	0.05	0.17	0.0993	0.10065	0.02448
Biological/Medical Sciences	0.11	0.22	0.1536	0.152188	0.02977

TABLE 4b. One-way ANOVA test results for reciter rate.

	SS	*df*	MS	*F*	Sig.
Between Groups	0.078	2	0.039	32.206	0.000
Within Groups	0.106	87	0.001		
Total	0.184	89			

Correlation Analysis Comparing Citer and Citation-based Measures

Our earlier study (Ajiferuke & Wolfram, 2010) revealed high correlations between citer and citation-based measures in most cases. Outcomes for the Social Sciences, Mathematical/Engineering, and Biological/Medical Sciences fields appear in Tables 6, 7, and 8, respectively. Note that all measures are significantly correlated based on Pearson's r and Spearman's rho tests (skewed distributions) for association at alpha levels of 0.05, except for one case. The only outcome that is not significant at an alpha level

DOI: 10.1002/asi

TABLE 5a. Descriptive statistics for ch-index.

	Min	Max	Mean	Median	St. dev.
Mathematical/Engineering Sciences	11	73	32.6333	31	17.537
Social Sciences	9	77	31.2333	27	16.962
Biological/Medical Sciences	33	180	93.5333	85.5	37.706

TABLE 5b. One-way ANOVA test results for ch-index.

	SS	*df*	MS	*F*	Sig.
Between Groups	75920.6	2	37960.3	56.46	0.000
Within Groups	58493.8	87	672.343		
Total	134414.4	89			

of 0.05 is the association between the number of citations per publication and ch-index values for the Social Sciences. This is similar to what was observed in our earlier study of LIS researchers, who can be categorized as social science researchers.

The high correlations indicate a close association between the citation and citer-based measures, but do not reflect changes in the ranking of outcomes. For each of the nine fields, the 10 authors were ranked based on their adjusted citation and citer counts. The differences in resulting ranks between the two measures were tabulated and grouped across the three broad subject areas (Table 9). The outcomes demonstrate that for the Social Sciences and Engineering/Mathematical Sciences more than half of the authors experience a change in rank, but this difference is no larger than two places. Note that any differences are bounded by the number of authors being ranked. Larger differences in rankings may be seen with a larger number of authors per field.

Differences in Citer and Citation-based Measures

The earlier results and analysis have revealed that citer and citation measures have significant correlations in most cases. However, these two measures can differ so that the number of citations may not reflect the reach of an individual article, or author by extension. Two scenarios are highlighted related to an increase in citations without new citers and differences

TABLE 6. Correlation coefficients between citer indices and other citation indices for Social Sciences scholars.

	Citation count	Adjusted citation count	Citations per publication	Adjusted citations per publication	h-Index	Adjusted h-index
Citer count	0.950 (0.000)*	0.935 (0.000)	0.611 (0.000)	0.561 (0.001)	0.869 (0.000)	0.884 (0.000)
Citers per publication	0.824 (0.000)	0.834 (0.000)	0.850 (0.000)	0.835 (0.000)	0.636 (0.000)	0.699 (0.000)
ch-index	0.839 (0.000)	0.813 (0.000)	0.367 (0.046)	0.318 (0.086)	0.976 (0.000)	0.964 (0.000)

*Significance level in parentheses.

TABLE 7. Correlation coefficients between citer indices and other citation indices for Mathematical/Engineering Sciences.

	Citation count	Adjusted citation count	Citations per publication	Adjusted citations per publication	h-Index	Adjusted h-index
Citer count	0.940 (0.000)*	0.940 (0.000)	0.808 (0.000)	0.818 (0.000)	0.932 (0.000)	0.942 (0.000)
Citers per publication	0.817 (0.000)	0.855 (0.000)	0.940 (0.000)	0.953 (0.000)	0.782 (0.000)	0.818 (0.000)
ch-index	0.897 (0.000)	0.858 (0.000)	0.625 (0.000)	0.640 (0.000)	0.973 (0.000)	0.969 (0.000)

*Significance level in parentheses.

TABLE 8. Correlation coefficients between citer indices and other citation indices for Biological/Medical Sciences.

	Citation count	Adjusted citation count	Citations per publication	Adjusted citations per publication	h-Index	Adjusted h-index
Citer count	0.943 (0.000)*	0.953 (0.000)	0.779 (0.000)	0.786 (0.000)	0.919 (0.000)	0.946 (0.000)
Citers per publication	0.711 (0.000)	0.748 (0.000)	0.956 (0.000)	0.938 (0.000)	0.668 (0.000)	0.682 (0.000)
ch-index	0.887 (0.000)	0.860 (0.000)	0.541 (0.002)	0.535 (0.002)	0.952 (0.000)	0.930 (0.000)

*Significance level in parentheses.

DOI: 10.1002/asi

TABLE 9. Change in author rank between adjusted citation count and citer count.

Change (adj. citation rank – citer rank)	Social sciences	Biological/ medical sciences	Engineering/ mathematical sciences
+2	2	0	2
+1	7	6	7
0	12	19	11
−1	7	4	9
−2	2	1	1

in citations and citers despite similar collaborative levels of citers.

An increase in citations does not necessarily result in new citers if the citations continue to come from the same set of citers. Conversely, new citers will always lead to gains in citations (but not necessarily gains in adjusted citations). Examples can be found where the number of citations continues to increase but the number of citers does not increase correspondingly, which indicates that citations might come from reciters that do not increase the reach of a publication. Based on a comparison of the data for the 40 authors (10 authors for four fields) from the extended Thomson Reuters dataset, most highly cited articles will gain adjusted citation as well as new citers. In addition, new citations are gained from reciters. Based on the cumulative data across the authors for articles that have been cited at least 100 times, a few dozen examples were found where the number of citers stays constant in a given year, while the cumulative adjusted citation count increases. The differences between these two values in a given year did not exceed four. These differences are small in comparison to the total number of citations. However, these differences become more apparent for lesser-cited articles where the cumulative number of citers plateaus while citation or adjusted citation counts may increase. Figures 3 and 4 demonstrate two cases where a gain in citations does not increase citers. In Figure 3, the article attracts six citations and three adjusted citations without attracting new citers over a 2-year period. In Figure 4, an article continues to attract new citations with minimal gains in the number of citers and adjusted citations. In both cases the increases in the numbers of citations do not reflect the stagnating reach of the article.

A comparison of articles can be made with roughly the same number of citations but different numbers of citers, and articles with roughly the same number of citers but different numbers of citations. As noted above, the number of citers is influenced by the average number of authors for citing articles. Two articles with same number of citations might have different numbers of citers only due to the fact that they have different numbers of collaborating authors in the citing articles. To control the effect of the average number of authors in citing articles, we compare articles with similar average numbers of authors in the citing articles. Recall that self-citations may still contribute to the number of citers if the cited author collaborates with someone who has not yet been counted as a citer. In this case, we should not exclude self-citations when determining the reach of the author because citing authors may be excluded simply because they have collaborated with the cited author.

In Table 10 the two articles have approximately the same number of citations (82 vs. 81) and levels of collaboration among citing authors, although different numbers of adjusted citations (i.e., 53 vs. 72) and quite different levels of reach (226 citers vs. 292 citers). The numbers of reciters are similar, with Article A having 48 reciters (226 − 178),

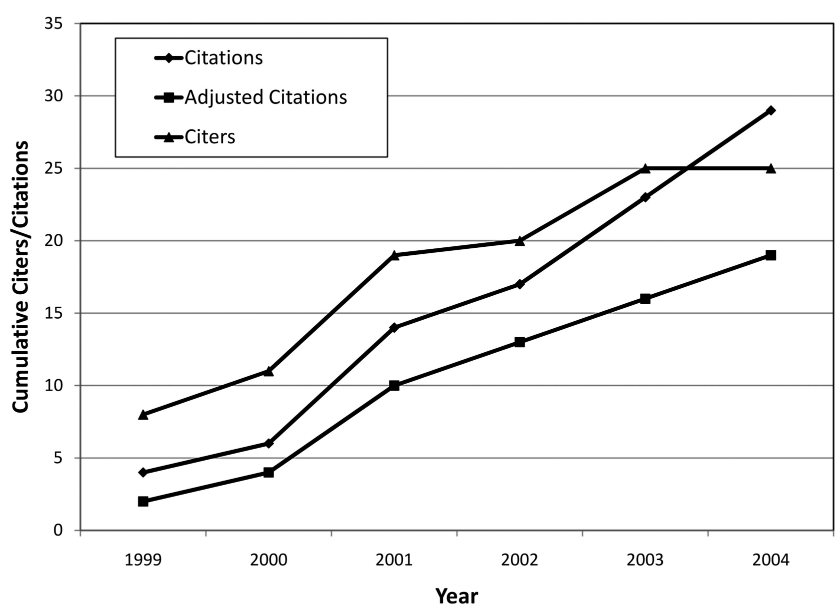

FIG. 3. Gaining adjusted citations without attracting new citers for a lesser-cited article.

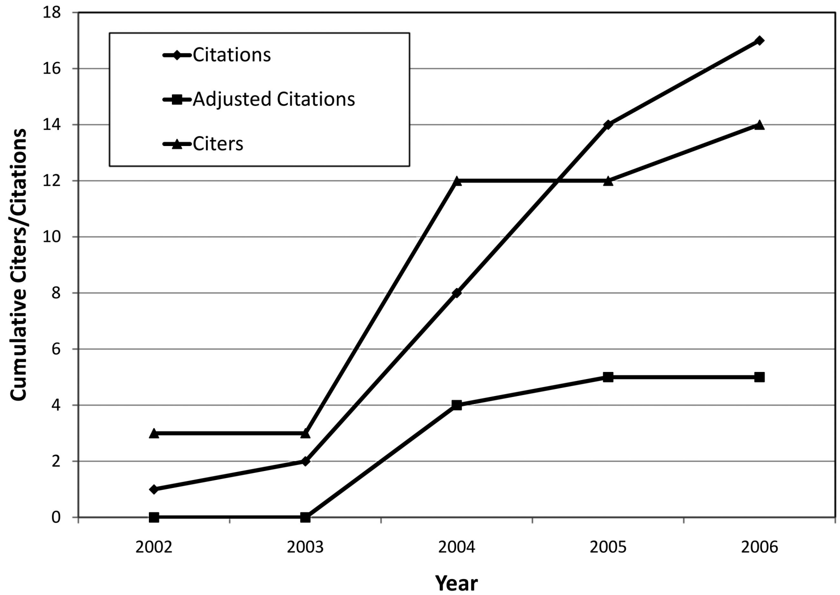

FIG. 4. Gaining citations without gaining adjusted citations or citers for a lesser-cited article.

TABLE 10. Articles with similar citation and collaboration levels, but different assessments of reach.

	Citations	Adjusted citations	Unique citers	Nonreciters	Avg # of authors per citing article
Article A	82	53	226	178	4.6
Article B	81	72	292	243	4.6

TABLE 11. Articles with similar reach and collaboration levels, but different citations.

	Citations	Adjusted citations	Unique citers	Nonreciters	Avg # of authors per citing article
Article C	442	401	1606	1295	4.97
Article D	392	361	1609	1410	4.97

and Article B having 49 reciters. These data tell us that although they receive a comparable number of citations, Article B outperforms Article A in terms of its reach based on the number of citers.

In Table 11 the two articles have similar levels of reach (1606 vs. 1609 citers) and levels of collaboration, but different numbers of citations (442 vs. 392). Article C has 311 reciters, whereas Article D has only 199 reciters. This example reveals that although Article C has more citations (and adjusted citations), it does not attract more citers than Article D because there are more reciters for Article C. These recitations add to the total citations but do not contribute to the reach of the article.

Discussion

Citer and citation-based measures produce largely similar outcomes, although they can provide complementary ways of assessing researcher impact or at higher levels, such as at the research group or institutional level. The present study focused on highly cited authors in those fields studied. However, even prolific, highly cited authors have less-cited articles. The above examples demonstrate that differences in citer and citation outcomes may also exist for lesser-cited works. Based on the current findings for the Social Sciences, in combination with the earlier pilot study involving library and information science (Ajiferuke & Wolfram, 2010), the Social Sciences may exhibit the largest differences between citer and citation-based measures. This is less the case with the Mathematical/Engineering Sciences and Biological/Medical Sciences studied, where the magnitudes of the differences between citer and citation measures may be large; however, the correlations are generally high. Reciter rates were also much lower for the Social Sciences, indicating that citers were less likely to continue citing the same work, contributing to lower adjusted citation counts. The lower correlations observed between the ch-index and citations per publication tabulations, in particular for the social sciences, suggest that the ch-index measure shares at least some of the limitations associated with the h-index (Costas & Bordons, 2007).

One could debate whether it is preferable to have a large number of citers who do not recite an author/work or if a large number of reciters is better, where recitation provides an indication of ongoing influence. Recitation itself could be studied from the perspective of recitation to the same work, or recitation to the same author, but different works. The second case could provide an indication of an ongoing following of cited author's research, or who the disciples are of a given author.

DOI: 10.1002/asi

The use of citer counts may be more "objective" than citation counts when levels of collaboration are not very high, as observed for the Social Sciences. Hyperauthorship can inflate citer counts to unreasonable levels and is more likely to occur in the life sciences, although if infrequently encountered it should not unduly influence outcomes. For prolific authors with high citation counts, the impact will be less noticeable than for less prolific authors with more modest citation counts. Citer counts for authors in fields where higher levels of coauthorship on citing articles are found will likely result in higher numbers of citers when compared with authors that have a comparable number of citations in fields where levels of coauthorship by citing articles are lower, and similar levels of recitation are observed.

The more fine-grained analysis comparing citations, adjusted citations, citers, and levels of recitation demonstrates that citation analysis on its own can exclude evidence that supports the extended reach of an article and its authors. Citation and adjusted citation counts can increase without contributing new citers, which do not extend the reach or influence of a given article. This was more evident with lesser-cited articles, where differences are relatively larger than for more highly cited articles. Such an analysis can have implications for assessing the merit of articles that are earlier in their life-cycle than citation classics, or for authors who are relatively new to the research landscape. In the case of the latter, an author's ch-index may be higher than another author with a comparable number of articles and citations.

Limitations of citer-based measures and the current study must be acknowledged. Because the number of citers is still based on the act of citation, some similar drawbacks exist as with citation measures. Classic challenges, such as the disambiguation of authors (homonyms) and variant author spellings still apply with citer analysis when assessing credit, both for those receiving credit and those providing credit. How one counts and provides credit will influence outcomes. The issue of recitation, for example, is addressed with the citer analysis by not providing additional credit to a citing author unless a different work by the cited author is involved. The level of collaboration exhibited on publications within a discipline can either benefit or disadvantage an author depending on the counting strategy used. Fractional counting acknowledges the distribution of credit, but creates an abstract level of recognition or attribution. Our focus has been on the number of citers, or authors who have been influenced by a given work. To say that one has been fractionally influenced by an idea or method seems arbitrary. Finally, citations (and citers) represent a uniformly applied, quantitative assessment of the influence of a work, but do not make qualitative assessments about influence.

Conclusions

In this study the authors explored the idea of citer analysis on a larger scale by examining citer and citation-based measures for several disciplines. The authors found citation counts, levels of collaboration by citing articles, recitation, and self-citation play a role in the number of citers an author's work attracts. The high correlations between citation and citer-based measures indicate a close relationship between the two, but they do result in somewhat different outcomes when authors are ranked based on adjusted citation or citer counts. The findings demonstrate that the number of unique citers can be used as a complementary measure to the adjusted number of citations to determine an article or author's influence or reach. Furthermore, citer analysis offers an alternative and complementary way to assess the impact or popularity of research by focusing on the reach of research, which may be less evident or even hidden if only citations are used. Citer data also offer a richer set of links among authors by which networks of relationships may be studied. The present study focused on highly cited researchers, but may be applied to researchers regardless of their citation history and can be applied at higher levels of aggregation, such as the research unit or institution level.

Future research will explore how citer and citation measures differ for authors who are not highly cited or through different levels of credit attribution to each citing author based on the citers' own citation or citer counts. The authors are currently investigating frequency distributions of recitation by citing authors based on numbers of different publications cited. Similar to the idea of citation image makers (White, 2000), this analysis will examine the distribution of authors who have been most influenced by the breadth of a cited author's work instead of the total number of citations received by a citing author across all publications.

Acknowledgments

This research was supported by a Thomson Reuters Citation Analysis Research Grant. We thank Thomson Reuters for additional data access and support. We thank Courtney Kennedy for providing research assistance and the anonymous reviewers for their helpful suggestions.

References

Ajiferuke, I., & Wolfram, D. (2010). Citer analysis as a measure of research impact: Library and information science as a case study. Scientometrics, 83(3), 623–638.

Bar-Ilan, J. (2006). An ego-centric citation analysis of the works of Michael O. Rabin based on multiple citation indexes. Information Processing & Management, 42(6), 1553–1566.

Bornmann, L., & Daniel, H.-D. (2008). What do citation counts measure? A review of studies on citing behavior. Journal of Documentation, 64(1), 45–80.

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.

Costas, R., & Bordons, M. (2007). The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level. Journal of Informetrics, 1(3), 193–203.

Costas, R., van Leeuwen, T.N., & Bordons, M. (2010). Self-citations at the meso and individual levels: Effects of different calculation methods. Scientometrics, 82(3), 517–537.

Cronin, B. (2001). Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices? Journal of

DOI: 10.1002/asi

the American Society for Information Science and Technology, 52(7), 558–569.
Cronin, B., & Overfelt, K. (1993). Citation-based auditing of academic performance. Journal of the American Society for Information Science, 45(2), 61–72.
Cronin, B., & Shaw, D. (2002). Identity-creators and image-makers: Using citation analysis and thick description to put authors in their place. Scientometrics, 54(1), 31–49.
Dieks, D., & Chang, H. (1976). Differences in impact of scientific publications: Some indices derived from a citation analysis. Social Studies of Science, 6(2), 247–267.
Garfield, E. (1955). Citation indexes for science: A new dimension in documentation through association of ideas. Science, 122(3159), 108–111.
Garfield, E. (1965). Can citation indexing be automated? In M.E. Stevens, V.E. Giuliano, & L.B. Heilprin (Eds.), Statistical association methods for mechanized documentation: Symposium proceedings (pp. 189–192). Washington, DC: National Bureau of Standards.
Garfield, E., & Sher, I.H. (1963). New factors in the evaluation of scientific literature through citation indexing. American Documentation, 14(3), 195–201.
Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences U S A, 102, 16569–16572.
Katsaros, D., Akritidis, L., & Bozanis, P. (2009). The f index: Quantifying of coterminal citations on scientists' ranking. Journal of the American Society for Information Science and Technology, 60(5), 1051–1056.
Kvam, P.H., & Vidakovic, B. (2007). Nonparametric statistics with applications to science and engineering. Hoboken, NJ: John Wiley & Sons.
Lindsey, D. (1989). Using citation counts as a measure of quality in science: Measuring what's measurable rather than what's valid. Scientometrics, 15(3–4), 189–203.
MacRoberts, M.H., & MacRoberts, B.R. (1989). Problems of citation analysis: A critical review. Journal of the American Society for Information Science, 40(5), 342–349.
MacRoberts, M.H., & MacRoberts, B.R. (2010). Problems of citation analysis: A study of uncited and seldom-cited influences. Journal of the American Society for Information Science, 61(1), 1–13.
Marshakova, I.V. (1973). A system of document connections based on references. Scientific and Technical Information Serial of VINITI, 6, 3–8.
Moed, H.F. (2005). Citation analysis in research evaluation. Dordrecht, Netherlands: Springer.
Phelan, T.J. (1989). A compendium of issues for citation analysis. Scientometrics, 45(1), 117–136.
Shapiro, F.R. (1992). Origins of bibliometrics, citation indexing, and citation analysis: The neglected legal literature. Journal of the American Society for Information Science, 43(5), 337–339.
Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. Journal of the American Society for Information Science, 24(4), 265–269.
Smith, L.C. (1981). Citation analysis. Library Trends, 30(1), 83–106.
Snyder, H., & Bonzi, S. (1998). Patterns of self-citation across disciplines (1980–1989). Journal of the American Society for Information Science, 48(6), 431–435.
Swanson, D.R. (1986). Undiscovered public knowledge. Library Quarterly, 56(2), 103–118.
Swarna, T., Kalyane, V.L., & Kumar, V. (2008). Homi Jehangir Bhabha: His collaborators, citation identity, and his citation image makers. Malaysian Journal of Library & Information Science, 13(2), 49–67.
White, H.D. (2000). Toward ego-centered citation analysis. In B. Cronin & Atkins H.B. (Eds.), The web of knowledge (ASIST Monograph Series, pp. 475–496). Medford, NJ: Information Today.
White, H.D. (2001). Authors as citers over time. Journal of the American Society for Information Science and Technology, 52(2), 87–108.
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(4), 327–355.
White, M.D., & Wang, P. (1997). A qualitative study of citing behavior: Contributions, criteria, and metalevel documentation concerns. Library Quarterly, 67(2), 122–154.

DOI: 10.1002/asi

Appendix

Sample author dataset with calculations

Publication #	Number of authors	Number of citations	Number of adjusted citations	Number of unique citers	Number of nonreciters	Reciter rate
1	8	140	130	809	643	=(809 − 643)/809 = 0.205
2	9	55	50	200	177	0.115
3	1	494	480	1212	1035	0.146
4	2	22	21	40	37	0.075
5	1	168	163	451	348	0.228
6	1	35	35	68	61	0.103
7	1	27	25	49	43	0.122
8	3	8	7	223	217	0.027
9	6	82	79	310	272	0.123
10	1	57	56	139	125	0.101
11	2	35	35	47	38	0.191
12	2	135	132	349	310	0.112
13	1	41	40	112	100	0.107
14	2	465	464	1313	1130	0.139
15	1	56	54	96	85	0.115
16	2	34	28	49	39	0.204
17	2	40	28	115	91	0.209
18	1	30	23	60	53	0.117
19	1	60	59	144	135	0.063
20	2	12	10	22	18	0.182
21	1	49	46	79	67	0.152
22	3	128	123	447	401	0.103
23	2	109	103	238	206	0.134
24	2	26	23	35	28	0.2
25	3	71	58	229	196	0.144

Adjusted citation count = (sum of the number of adjusted citations column) = 2272.
Adjusted citations per publication = 2272/25 = 90.88.
Citer count = (sum of the number of unique citers column) = 6836.
Citer per publication = 6836/25 = 273.44.
h-index (based on number of citations column) = 22.
Adjusted h-index (based on number of adjusted citations column) = 22.
ch-index (based on number of unique citers column) = 24.
Average reciter rate = (average of reciter rate column) = 0.137.