Article

Journal of Contemporary Criminal Justice
2023, Vol. 39(3) 327–340

DOI: 10.1177/10439862231170972
journals.sagepub.com/home/ccj
Sage

Citation Data and Analysis: Limitations and Shortcomings

John L. Worrall[1] and Ellen G. Cohn[2]

Abstract

As a means for measuring scholarly influence, citation analysis has several limitations and shortcomings. We first review the main sources of citation data (Web of Science, Scopus, Google Scholar, and information collected directly from reference lists) and discuss the shortcomings of each source. Next, we review five significant limitations of citation analysis as a methodology (academic over popular interest, various motivations for citing, manipulation potential, failure to account for author ordering, and citations only appearing in "indexed" journals). The issues we touch on set the stage for the remainder of the articles in this special issue.

Keywords

citation, influence, Web of Science, Scopus, Google Scholar

Citation analysis has long been employed as a method for determining scholars', journals', and university departments' impact on the field. It has also been used to gauge the influence of specific works (journal articles, books, etc.) on later research products, as researchers tend to cite works that they find impactful and valuable. There are two primary assumptions made by scholars using citation analysis: that the most-cited works are those that were most important to the research and that citations indicate influence (Meadows, 1974). The primary advantage of citation analysis is that it is objective, quantitative, and replicable, and is less likely to be influenced by personal bias or special interest than other measures of measuring scholarly influence and prestige, such as peer rankings and prestige surveys or the receipt of scholarly awards and prizes.

[1]The University of Texas at Dallas, Richardson, USA
[2]Florida International University, Miami, USA

Corresponding Author:
John L. Worrall, Criminology and Criminal Justice, The University of Texas at Dallas, Richardson, TX 75080, USA.
Email: worrall@utdallas.edu

Counting journal publications by individuals or departmental faculty has also been used to assess prestige. However, while publication counts are more quantitative and objective than peer rankings, they do not indicate the impact of publications; instead, they provide a measure of faculty or institutional productivity. The publication of an article is no guarantee that it will be read or referred to by other scholars. In fact, it has long been known that the distribution of citations is skewed, so that a small number of publications are highly cited while the majority receive few or no citations (see, e.g., Zhang et al., 2017).

As a means for measuring scholarly influence, citation analysis does have some problems. In this article, we discuss several of them. First, we review the main sources of citation data and discuss the shortcomings of each source. Second, we review five significant limitations of citation analysis as a methodology. The issues we touch on set the stage for the remainder of the articles in this special issue.

Citation Data Limitations

There are four primary sources of citation data that are used by researchers examining the field of criminology and criminal justice. Three of these are available online. The first is the Web of Science, which is owned by Clarivate, a public analytics company. The second is Scopus, which is operated by Elsevier, an academic publishing company based in the Netherlands. The third is Google Scholar, a web search engine owned by Google. Finally, the fourth source of citation data involves collecting information directly from reference lists of selected journals (or other scholarly sources).

The Web of Science

The Web of Science is a fee-based platform that indexes scholarly literature. The Core Collection comprises six main databases. Of these, the Social Sciences Citation Index includes most journals relevant to criminology. In addition, there are several regional citation indexes, which focus on specific countries or regions of the world. The Web of Science Core Collection covers more than 250 disciplines and includes about 1.5 billion references from thousands of peer-reviewed journals, scholarly books, research articles, editorials, and so on (Clarivate, 2022b).

Clarivate also produces the *Journal Citation Reports* (JCR) each year, based on citations from more than 21,000 journals indexed in two key Web of Science Indexes—the Social Sciences Citation Index and the Science Citation Index. The JCR is a journal-level resource, rather than a source of information on citations of individual works. For each journal, the JCR provides a considerable variety of information, including the journal impact factor and the impact factor trend going back to 1997 if the journal has been in print that long (Clarivate, 2022a). The journal impact factor is frequently used by scholars as a measure of a journal's relative importance; some researchers conducting citation analysis use the impact factor as a metric when deciding which journals to study (see, e.g., Moeller, 2019).

While the Web of Science provides citation data from an extremely large number of journals in a wide variety of disciplines, it has some significant limitations. First, the list of journals included in the Web of Science is not stable. It changes constantly as new journals are added and others are removed from the master journal list. In the past, the Web of Science provided a list of these changes (in 2012, the list was 20 single-spaced pages long, with more than 40 journals listed per page; see Cohn & Farrington, 2012), but this information is no longer provided on the website. This instability affects the accuracy of longitudinal research. In addition, some issues of a journal or book series may be omitted. In their study of citations in *Crime and Justice: A Review of Research*, Cohn et al. (2018) found that the Web of Science did not index 15 issues of the book series.

Another serious concern is that any errors present in the source reference lists, such as spelling errors, errors in publication dates, and incorrect initials, are carried over to the Web of Science. As Sweetland (1989) pointed out, ". . . the assumption is that the citations are accurate. Unfortunately, considerable evidence suggests that such an assumption is questionable" (p. 292). Spelling errors and incorrect first initials in reference lists are surprisingly common. For example, Cohn and Farrington (1996) found Travis Hirschi's works listed under names such as Hirsch, Hirsh, and Hirshi and with first initials of "P" and "L" instead of the correct "T. Hirschi." A related concern is that some journals use the generic "et al." in reference lists, rather than listing the names of all the authors of the work. This means that those additional authors will not receive credit for the citation in the Web of Science. Some journals, such as *Journal of Criminal Justice*, do this rather frequently. The first volume published in 2021 (Volume 72) contained 20 articles; of these, 13 included at least one reference that omitted one or more authors. One article alone (Seo et al., 2021) contained nine such references that together omitted a combined 21 authors.

The dates of works in reference lists may also be listed incorrectly, and these errors will also be brought over to the Web of Science. Sweetland (1989) found errors in 28% of cited articles and estimated that about 7% of these errors were serious enough to make it difficult to obtain a work based on the reference information. The errors he identified included misspelled surnames and incorrect author initials as well as omitted or incorrect journal titles, page numbers, volume numbers, dates, and so forth. Moed (2002) found author errors in about 7% of cited articles in the Science Citation Index (one of the primary databases within the Web of Science). Siebers and Holt (2000) examined five leading medical journals and found reference error rates ranging from 4.1% to 40.3%. Other studies of citation accuracy have reached similar conclusions (see, e.g., Buchanan, 2006; Lok et al., 2001; Wyles, 2004). It is, however, important to recognize that these errors occur in the primary documents indexed by the Web of Science; they are not made by the database itself.

Because many reference lists only include an author's last name and initial(s), citations to multiple scholars with the same surname and initials may be merged. This problem is exacerbated by the fact there are scholars even within the same field who may have the same full name (such as the multiple Richard Berks, David Smiths, and Tony Wards) or who use middle initials that may be omitted in reference lists (e.g., Ellen G. and Ellen S. Cohn). A similar problem is that some references use a "nickname" first initial (Chapman, 1989).

The Web of Science does provide algorithmically created author records, which are "groups of publications that are likely by one person" (Clarivate 2019). However, a publication may be assigned to the wrong author record or to multiple records; in addition, the same author may have several author records. Chapman (1989) also pointed out that married female scholars may publish under more than one surname (e.g., Ilene H. Nagel and Ilene Nagel Bernstein). This problem may extend beyond female scholars to any individual who changes their surname.

Finally, the Web of Science databases include self-citations. While these are perfectly justifiable, as many authors build on their own prior research, they should be excluded in research that focuses on scholarly influence on the field.

Scopus

The second source of data is Scopus, a citation and abstract database which has been produced by Elsevier since 2004. As of 2020, Scopus included content from more than 25,100 active titles and more than 5,000 publishers. This included more than 23,000 peer-reviewed journals, almost 300 trade publications, more than 850 book series, almost 10 million conference papers, and more than 44 million patent records, as well as articles in press from more than 5,500 journals The Scopus books title list includes more than 210,000 books. In total, the database contains almost 78 million records (Scopus, 2020).

While Scopus is very user-friendly, it also has several limitations (see, e.g., Dess, 2006). Like the Web of Science, the list of publications covered is not static; journals may be added to or removed from the list of sources covered. Another concern is that the database primarily focuses on journals; while book series may also be included and book coverage in Scopus is expanding, the database is primarily focused on citations from journals (Scopus, 2020). In addition, although the oldest record in Scopus dates to 1788, cited references are not included in documents published prior to 1970 (Scopus, 2020). This means that citation tracking is extremely limited.

Dess (2006) discovered that the results of a search will vary depending on the author in which search terms are entered, particularly if the "search within" function is used. A recent search we conducted employing the same keywords used by Dess found the same problem to exist. Dess also found that retrievals of web sources and patents varied greatly over time (e.g., he found that over the course of several days, patent searches ranged from 470 hits to only 4).

Scopus is also heavily biased toward publications from North America and Western Europe. In 2020, there were more than 6,600 active titles from North America and 12,170 from Western Europe; this may be compared with 2,230 from Asia Pacific, 790 from Latin America, and 1,750 from East Europe and Russia (Scopus, 2020). In addition, Scopus is biased toward the natural sciences (life and health sciences), while there is more limited coverage of the social sciences. As of 2020, there were a total of 14,448 titles in the health sciences and another 7,295 in life sciences. While there are a total of 12,464 titles in the social sciences, this is misleading as this category also includes more than 4,700 titles in the arts and humanities.

Finally, both Scopus and Web of Science are subscription-based platforms; to conduct searches, it is necessary to purchase a subscription or have access to a library that has purchased one. University libraries often must decide whether to choose one of the two platforms or to bear the expense of subscribing to both

Google Scholar

Google Scholar, a free Internet search engine that indexes peer-reviewed journals, books, and other scholarly literature, as well as selected "scholarly web pages," is the third source of citation data. Like Scopus, Google Scholar has been in operation since 2004. It uses an automated web crawler to trawl full-text journal content and bibliographic databases. A major benefit to Google Scholar is that it is free, providing access to information that may not otherwise be available to individuals who cannot access fee-based services such as the Web of Science or Scopus. Another positive feature is that it includes more non-journal materials than the other sources, and indexes documents in a wide variety of languages, including those that do not use the Latin alphabet (e.g., Japanese, Russian, Chinese, Hebrew, and Arabic) (Bauer & Bakkalbasi, 2005).

Google Scholar also has a number of serious limitations, and several scholars (see, e.g., Jascó, 2008; Neuhaus & Daniel, 2008) have expressed concern over its use as a tool for citation analysis. One major concern is that the site provides little or no information about the coverage so researchers cannot determine which journals are trawled, how complete the site's coverage is, or how often Google Scholar is updated. As Jascó (2008) pointed out, "From the launch of the service, it has been hopeless to derive any factual information from Google, Inc. regarding the dimension of the content of the database, its size, girth (width, length, and depth combined), or the sources included" (p. 440). In fact, the Google Scholar's (2022a) online "help" page answers the question of what specific journals are covered by stating, "Ahem, we index papers, not journals." This is particularly a concern given that Google Scholar appears to index predatory and peer-reviewed journals (Kolata, 2017).

There are also some concerns about the functioning of the search engine. For example, when presenting the results of a search, Google Scholar defaults to sorting by relevance. While there is an option to sort by date, selecting this only presents articles added in the last year. As a test, we searched for works on "terrorism" in Google Scholar and were presented with about 2,050,000 results, sorted by relevance. When we sorted these works by date, we were only given about 8,160 results. It is also not possible to limit a search (other than a search of case law) to specific types of works, such as journal articles or books; the only option seems to be to limit a search to "review articles" but this term is not clearly defined.

Journal Reference Lists

The final source of citation data involves directly collecting information from the reference lists of journals, scholarly books, textbooks, etc., and counting the number of citations of a given scholar, work, or journal. This method was pioneered by Cohn and

Farrington (1990) and has been used by them in a series of studies more than 35 years (see, e.g., Cohn et al., 2023). It avoids many of the problems and limitations found when using online databases or search engines. It allows for replicability as the raw data are readily available to any researcher who has access to the journals being studied. It also permits researchers to focus on a specific field or subfield, allows them to obtain and include the names of additional co-authors that may be hidden within an "et al.," and enables the identification (and exclusion if desired) of self-citations.

Even this source has limitations, however. First, it is extremely labor-intensive and time-consuming, as the data are obtained directly from the reference lists of each journal being examined. Although this information is now available online and the use of Excel permits automation of much of the process, it still requires much manual effort on the part of the researcher. Second, because of the difficulty of the data collection process, research typically must be restricted to citations in a relatively small number of journals. This means that the influence of scholars who do not publish in the selected journals (e.g., a criminologist who publishes primarily in journals focusing on a cognate discipline such as psychology or sociology) may be underestimated.

Citation Analysis Shortcomings

It is clear by now that the data used for citation analysis are problematic on several fronts. They are prone to author and database errors, plus citation analysis relies solely on "indexed" journals, so a potentially large number of journals are excluded. Data limitations notwithstanding, online sources of citation data are some of the best we have—short of poring through journals and tallying up citations by hand, which is of course prone to its own set of errors. Even if citation data were completely "clean," however, citation analysis itself is beset by several shortcomings. Indeed, the shortcomings are the impetus for this special issue! (We leave it to readers to decide whether the alternatives improve over the originals.)

Citation analysis faces at least five significant shortcomings. First, it is limited to academic, not popular interest (in fairness, however, that may not be its intent). Second, quality is not necessarily the key consideration, as there are many, potentially conflicting motivations for citing scholarship. Third, "manipulation" is a common problem. Fourth, citation analysis does not count for author ordering; a distant fifth author, for example, typically gets the same credit as the lead author. Finally, one of the data problems covered earlier, that is, something less than the population of journals being indexed, rears its head here; nonindexed journals do not factor into citation analysis (unless manually counted), leaving a large swath of scholarship in a sort of intellectual purgatory where authors do not get (citation-based) credit for their professional accomplishments.

Academic Versus Popular Interest

Critics of citation analysis claim that it is limited to academic, not popular interest. There are two responses to this critique. On one hand, there is no intent for traditional citation metrics to gauge the popular interest; they are not designed for that. In other words, citations are intended for scholars to recognize other scholars for their intellectual contributions to the field and current research. On the other hand, a great deal of academic research has concrete real-world implications (e.g., cures for disease and successful crime reduction policies). Ignoring these can fuel the conception of an insular, self-interested ivory tower community of experts who care for little else than sharing ideas among themselves.

Whether citations should be focused inward toward the scholarly community or outward toward society at large hinges one's definition of scientific value. In the first view, when a scientist cites a paper, it is done to display how useful or relevant a particular publication was to inform and shape the present research. Under this conception, the most frequently cited articles are the most useful for advancing science. Increasingly, though, scientists are being pushed to enhance the visibility of their research beyond the scientific community, especially when it comes to taxpayer funding of scholarly research (e.g., Alla et al., 2017; Langfeldt & Scordato, 2015). This takes us from a focus solely on scientific relevance to one of societal relevance. As the second section of this special issue attests, many of the more novel methods being developed are primarily concerned with questions of societal rather than scholarly influence.

Quality Concerns

A citation is a citation. Or is it? Citations can be favorable, or they can be critical. A scholar might just as easily praise a predecessor as the inspiration for a current project or tear the same person down by critiquing arguments and methodologies. What, then, is the theoretical meaning of a citation? How should one be interpreted? Two early pioneers in citation analysis, the Cole and Cole (1973), wrote in that "The number of citations is taken to represent the relative scientific significance or 'quality' of papers." That assumption persists, with many scholars submitting that citations measure quality more than anything else (e.g., Abramo & D'Angelo, 2011; Durieux & Gevenois, 2010), and in criminal justice/criminology, it appears that the *majority* of citations are generally positive or neutral, rather than critical (see Cohn & Farrington, 1994).

Nevertheless, research in this area has revealed the complex, multifaceted nature of citation behavior. L. C. Smith (1981) wrote that "Not enough is known about the 'citation behavior' of authors—why the author makes citations, why he makes his particular citations, and how they reflect or do not reflect his actual research and use of the literature" (p. 99). His concerns sparked a surge in studies of referencing behavior (e.g., Bornmann & Daniel, 2008; Nicolaisen, 2007; Thornley et al., 2015). From this literature, two competing perspectives have emerged. One describes citations as an intellectual exercise, and the other describes "citing" as a social process (e.g., to recognize one's friends and network acquaintances). To the extent, citations are social; then, they are not primarily concerned with academic quality (Aksnes, 2005).

Another issue concerns the *role* of references in a scientific paper. As early as 1964, Garfield identified 15 separate reasons why people cite other scholarship in their work (reprinted in Garfield, 1977). They include providing background reading, referencing methods, acknowledging the work of pioneers in the area, supporting claims, and disputing others' claims, among many others. More recently, Small (1982) identified five functions of citation: refuting, noting, reviewing, applying, and supporting. None of these functions are expressly concerned with quality. Within any one article, the "textual functions" (Aksnes et al., 2019) of referencing serve multiple goals, from acknowledging seminal works to merely providing background material.

Studies have also examined the incentive structure for citations. Look no further than self-citations. Creating visibility for one's work may superseded quality concerns. Of course, this extends to programs, departments, universities, journals, and pretty much every other dimension of academia; incentive structures recognize publications and thus citations. Accordingly, some have described citations as tools of persuasion: "To persuade the scientific community of the value and importance of their publication, authors are using references as rhetorical tools" (Aksnes et al., 2019, p. 5).

With growing disagreement over what citations represent, studies have attempted to validate them as quality metrics by comparing them to peer-reviewed outcomes. The methods are varied and beyond the scope of this article, but most studies have found a "moderately positive" correspondence between citation volume and peer assessments (Aksnes et al., 2019, p. 6). In other words, citations *reflect* quality, but they are not solely indicative of it. Combining this with the many reasons for citations, quality is but one of many distinct considerations. In addition, we have not even touched on the meaning of quality, a concept with its own distinct dimensions (Polanyi, 1962).

Manipulation

Reflecting back on some of his more highly cited works, physiological psychologist Robert Douglas (1992) authored an illuminating commentary in *Psychological Bulletin*, "How to write a highly cited article without even trying." His prescription for success is as follows:

> If you were to set out in cold blood to write a highly cited article, your best bet would be to devise or revise a paper-and-pencil test of personality or motivation, improve on a commonly-used method, coin a snappy new word or phrase, or think of a new way to apply statistics. (p. 405)

Extrapolating to criminology and criminal justice, his points resonate. For example, Alex Piquero's most-cited article, "Using the correct statistical test for the equality of regression coefficients" (Paternoster et al., 1998), did not tackle a substantive criminological question. Similarly, the most-cited criminological theorists tend to be those connected with the development or proposal of a new or fresh paradigm to explain criminal behavior (see, e.g., Table 5, p. 213, in Cohn & Farrington, 1994).

Douglas also pointed out that when a scholar attempts to "make sense" of a timely issue, they can boost their citation potential. "Making sense" is the province of review essays, literature syntheses, and meta-analyses, and if timed appropriately (e.g., first to print before anyone else authoring similar scholarship) and executed well, they can indeed be cited extensively. For example, the criminologist Travis Pratt's top four most-cited studies are all meta-analyses, according to Google Scholar (e.g., Pratt & Cullen, 2000).

It seems, then, that there are critical ingredients in the citation recipe. But what role does intent play? On one hand, Douglas (1992) attributed his citation success more to "improbable events" (p. 407) than grand design. On the other hand, "coldhearted authorship" (Douglas, 1992, p. 407), or for that matter editorial policy, may be capable of generating frequent citations. Even author laziness can boost citations. In a fascinating study, two University of California Los Angeles electrical engineers (Simkin & Roychowdhury, 2003) were able to demonstrate mathematically that scholars' copying of references from other works can account for a good deal of the citation dynamics in scientific papers. Specifically, they found that if a hypothetical scientist picks three papers at random, cites them, and copies a quarter of the references from each, their actions will account for the citation distribution. In their words, "[s]imple mathematical probability, not genius, can explain why some papers are cited a lot more than the other."

It is also clear that editorial policy can drive citation patterns. Researchers have confirmed this across the sciences. For example, some journals may require, as a condition of acceptance, authors to cite works in the same journal to enhance its impact factor (e.g., Miller, 2002; R. Smith, 1997). Alternatively, authors may themselves engage in the practice of citing papers from the journal they are submitting to, even if the research has little bearing on their own, which can then influence impact (e.g., Yang et al., 2016). Researchers have even discovered "citation cartels," wherein authors from, say, two journals have an inordinate number of citations to a third journal, to boost the latter's impact factor (e.g., Davis, 2012). Finally, editorial publications within journals can inflate impact factors when they excessively refer publications in the same journal (e.g., Heneberg, 2014).

Not long ago, criminology was rocked by scandal involving the latter strategy. Thomas Baker (2015) found that the *Journal of Criminal Justice* (JCJ) shot to number one in the (then Thompson Reuters) impact factor rankings, despite consistently ranking in the second or third quartiles of criminology and penology journals in previous years. He wrote, "Surely this meteoric rise deserves further investigation as it can provide insight into how *JCJ* became the top-ranked journal in the field . . ." (p. 5). He found that articles and editorials authored by the editor contributed substantially to *JCJ*'s increased prominence. His study even caught the attention of a writer for the *Chronicle of Higher Education*. Bartlett (2015) noted one particularly egregious example: "In the most eyebrow-raising instance, one four-paragraph editorial, published in 2014, didn't take up even a single page yet managed to have 47 citations, all to the *Journal of Criminal Justice*."

In a spirited response to Baker (and presumably to Bartlett), the *JCJ* editor claimed that impact factor manipulation was not his intent:

> The reasons for writing [the editorials] are varied but include the desire to influence researchers, to suggest calls for research, to comment on policy issues, to promote the work of scholars publishing in the journal, to promote the work of editorial board members, and to serve as a provocateur on controversial issues. (DeLisi, 2015, p. 5)

Regardless of whether the accuser or the accused was correct, *JCJ* has maintained a fairly high ranking since 2015. In 2021, for example, it was ranked sixth of 69 journals by Clarivate in its impact factor rankings for the criminology and penology category. That same year, though, it was surpassed by another historically lower-ranking journal, the *American Journal of Criminal Justice*, but its sudden rise in the rankings is most likely attributable to a high-profile special issue devoted to the COVID-19 pandemic and crime issues.

Multi-Authorship

Standard citation counts fail to accommodate the varied contributions involved in multi-authorship. This is problematic because the average number of authors on scientific papers has been steadily increasing over time (Kennedy, 2003), including in criminal justice and criminology (Fisher et al., 1996; Tewksbury & Mustaine, 2011).

It is difficult to weigh the contribution of each author in a multi-authored paper. Should the last author be given the same credit as the first? There is no easy answer because there are no standards for reporting author order. The convention is to list authors in descending order of their contribution, with the lead author getting the most credit because the study was their brainchild. Sometimes, though, authors list themselves alphabetically, and in some instances, they include notes to the effect that all authors contributed equally. In addition, in the case of doctoral students and their mentors, the former sometimes give the latter lead authorship because it looks good for the students' job prospects, even though the ideas came from the mentors. With no standard for reporting author order, it becomes nearly impossible to gauge contributions, calling citations' value into question. Under the current scheme in our field, a scholar who makes it a practice of being the last or nearly last author on a huge number of publications (sometimes called "gift authorship") gets potentially the same credit as a sole author of significantly fewer publications.

For some years now, information scientists have proposed alternative methods to evaluate individual scholars' contributions to multi-authored scholarship (e.g., Egghe, 2008; Galam, 2011; Schreiber, 2009). Many are variants on the familiar *h*-index (e.g., Salman et al., 2021), and they fractionally allocate authorship credit based on specific criteria (e.g., Vavrycuk, 2018). Most recently, there is a push to account for "scientific leadership" in author ordering (Hirsch, 2019; Simoes & Crespo, 2022), but the fruits of those labors have yet to catch on throughout the social sciences, and particularly in criminal justice/criminology, where it remains a tradition or convention to treat all citations with equal weights. In addition, besides, any formula or algorithm that purports to distinguish leaders from followers cannot possibly achieve its goal without knowing and understanding scholars' reasons for their author ordering schemes.

Journal Indexing Problems

Clarivate's curation process for the Web of Science relies on 28 criteria to evaluate journals for potential indexing (Clarivate, 2022c). In the first phase, called initial triage, journals are scrutinized to ensure full contact information is available (titles, publishers, journal URLs, etc.). Next, editorial triage is designed to ensure, among other things, that the journal contains scholarly content, uses clear language, and is timely. The third phase, editorial evaluation, is where editors check for ". . . alignment between the journal's title, stated scope, the composition of its editorial board and authorship, and its published content" (Clarivate, 2022c). The final step assesses impact. This phase is designed to identify the most influential journals in the field, using citation activity as a key selection criterion. In addition, "The content in the journal should be of interest, importance, and value to its intended readership and to Web of Science subscribers" (Clarivate, 2022c). Elsevier's Scopus employs similar selection criteria, including most notably quality (scientific work), diversity (diversity of authorship and in editorial boards), and relevancy (appeal to international user base).

Clarivate and Elsevier get credit for clearly articulating the criteria they rely on to decide whether a journal will be indexed, but even with multiple areas to consider, the decision-making process remains relatively opaque (e.g., submitting journals do not participate in the Content Selection and Advisory Board meetings for Elsevier). Journals that are rejected for indexing *do* get feedback, and they can be considered again, but the whole of the process is not completely transparent. It is even less transparent for Google Scholar, which has become particularly popular because it is freely available. Google Scholar is a "crawler-based" search engine, not a list of indexed publications. It relies on "spiders" or "bots," which are basically automated programs that scour the web to identify content worthy of inclusion. Once content is identified, Google "invites" content from trusted sources to be considered. As Google states on its website, "we work with publishers of scholarly information to index peer-reviewed papers, theses, preprints, abstracts, and technical reports from all disciplines of research and make them searchable . . ." (Google Scholar, 2022b). Importantly, Google Scholar is limited to original research and technical reports that are either made freely available in full or contain the full author-written abstract in the case of peer-reviewed articles located behind paywalls.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Abramo, G., & D'Angelo, C. A. (2011). Evaluating research: From informed peer review to bibliometrics. *Scientometrics*, *87*, 499–514.

Aksnes, D. W. (2005). *Citations and their use as indicators in science policy: Studies of validity and applicability issues with a particular focus on highly cited papers* [Doctoral thesis, University of Twente].

Aksnes, D. W., Langfeldt, L., & Wouters, P. (2019). Citations, citation indicators, and research quality: An overview of basic concepts and theories. *SAGE Open*, *9*(1),1–17. https://doi.org/10.1177/2158244019829575

Alla, K., Hall, W. D., Whiteford, H. A., Head, B. W., & Meurk, C. S. (2017). How do we define the policy impact of public health research? A systematic review. *Health Research Policy and Systems*, *15*, Article 84.

Baker, T. (2015). An evaluation of journal impact factors: A case study of the top three journals ranked in criminology and penology. *The Criminologist*, *40*(5), 5–10.

Bartlett, T. (2015, September 23). The journal that couldn't stop citing itself. *The Chronicle of Higher Education*. https://www.chronicle.com/article/the-journal-that-couldnt-stop-citing-itself/

Bauer, K., & Bakkalbasi, N. (2005). An examination of citation counts in a new scholarly communication environment. *D-Lib Magazine*, *11*(9). http://dlib.org/dlib/september05/bauer/09bauer.html

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

Buchanan, R. A. (2006). Accuracy of cited references: The role of citation databases. *College and Research Libraries*, *64*, 292–303.

Chapman, A. J. (1989). Assessing research: Citation-count shortcomings. *The Psychologist*, *8*, 336–344.

Clarivate. (2019). *Web of Science core collection: Author search beta*. http://images.mail.discover.clarivate.com/Web/ClarivateAnalytics/%7B93be44b6-6532-4194-8816-7f9d9e3f09a7%7D_WoS_CC_PROFILES_QRG.pdf?utm_campaign=EM_Profiles_Researchers_SAR_Global_2019&utm_medium=email&utm_source=Eloqua

Clarivate. (2022a). *Journal citation reports*. https://jcr.clarivate.com/jcr/home

Clarivate. (2022b). *Web of Science core collection*. https://clarivate.com/webofsciencegroup/solutions/web-of-science-core-collection/

Clarivate. (2022c). *Web of Science journal evaluation process and selection criteria*. https://clarivate.com/products/scientific-and-academic-research/research-discovery-and-workflow-solutions/web-of-science/core-collection/editorial-selection-process/editorial-selection-process/

Cohn, E. G., & Farrington, D. P. (1990). Differences between British and American criminology: An analysis of citations. *British Journal of Criminology*, *30*, 467–482.

Cohn, E. G., & Farrington, D. P. (1994). Who are the most influential criminologists in the English-speaking world? *British Journal of Criminology*, *34*(2), 204–225.

Cohn, E. G., & Farrington, D. P. (1996). Crime and Justice and the criminal justice and criminology literature. *Crime and Justice: A Review of Research*, *20*, 265–300.

Cohn, E. G., & Farrington, D. P. (2012). *Scholarly influence in criminology and criminal justice*. Nova Science Publishers.

Cohn, E. G., Farrington, D. P., & Skinner, G. (2023). *Most influential scholars in criminology and criminal justice, 1986-2020*. Springer.

Cohn, E. G., Iratzoqui, A., Farrington, D. P., Piquero, A. R., & Powell, Z. A. (2018). Most-cited articles and authors in Crime and Justice, 1979-2015. *Crime and Justice: A Review of Research*, *47*, 475–508.

Cole, J. R., & Cole, S. (1973). *Social stratification in science*. The University of Chicago Press.

Davis, P. (2012). The emergence of a citation cartel. *The Scholarly Kitchen*, *10*, 15–17.

DeLisi, M. (2015). Response to "an evaluation of journal impact factors: A case study of the top three journals ranked in criminology and penology." *The Criminologist*, *40*, 5–6.

Dess, H. M. (2006). Database reviews and reports: Scopus. Issues in Science and Technology Librarianship, *45*. http://www.istl.org/06-winter/databases4.html?ref=Sex%C5%9Ehop.Com

Douglas, R. J. (1992). How to write a highly cited article without even trying. *Psychological Bulletin*, *112*, 405–408.

Durieux, V., & Gevenois, P. A. (2010). Bibliometric indicators: Quality measurements of scientific publication. *Radiology*, *255*, 342–351.

Egghe, L. (2008). Mathematical theory of the h- and g-index in case of fractional counting of authorship. *Journal of the American Society for Information Science and Technology*, *59*, 1608–1616.

Fisher, B. S., Vander Ven, T. M., Cobane, C. T., Cullen, F. T., & Williams, N. (1996). Trends in multiple authored articles in criminology and criminal justice: A comparative analysis. *Journal of Criminal Justice Education*, *9*, 19–38.

Galam, S. (2011). Tailor based allocations for multiple authorship: A fractional gh-index. *Scientometrics*, *89*, 365–379.

Garfield, E. (1977). Can citation indexing be automated? In E. Garfield (Ed.), *Essay of an information scientist* (Vol. 1, pp. 84–90). ISI Press.

Google Scholar. (2022a). *Google Scholar search help*. http://scholar.google.com/intl/en/scholar/help.html

Google Scholar. (2022b). *Publisher support*. https://scholar.google.com/intl/en/scholar/publishers.html

Heneberg, P. (2014). Parallel worlds of citable documents and others: Inflated commissioned opinion articles enhance scientometric indicators. *Journal of the Association for Information Science and Technology*, *65*(3), 635–643.

Hirsch, J. E. (2019). Hα: An index to quantify an individual's scientific leadership. *Scientometrics*, *118*, 673–686.

Jascó, P. (2008). The pros and cons of computing the h-index using Google Scholar. *Online Information Review*, *32*, 437–452.

Kennedy, D. (2003). Multiple authors, multiple problems. *Science*, *301*, Article 733.

Kolata, G. (2017, October 30). Many academics are eager to publish in worthless journals. *The New York Times*, D5.

Langfeldt, L., & Scordato, L. (2015). *Assessing the broader impacts of research: A review of methods and practices* (NIFU Working Paper 8/2015). https://nifu.brage.unit.no/nifu-xmlui/bitstream/handle/11250/282742/NIFUworkingpaper2015-8.pdf?sequence=1&isAllowed=y

Lok, C. K. W., Chan, M. T. V., & Martinson, I. M. (2001). Risk factors for citation errors in peer-reviewed nursing journals. *Journal of Advanced Nursing*, *34*, 223–229.

Meadows, A. J. (1974). *Communication in science*. Butterworths.

Miller, J. B. (2002). Impact factors and publishing research. *Scientist*, *16*(18), 11.

Moed, H. F. (2002). The impact-factors debate: The ISI's uses and limits. *Nature*, *415*, 731–732.

Moeller, K. (2019). Ranking Nordic criminologists by impact and prestige. *Journal of Criminal Justice Education*, *30*, 536–550.

Neuhaus, C., & Daniel, H.-D. (2008). Data sources for performing citation analysis: An overview. *Journal of Documentation*, *64*, 193–210.

Nicolaisen, J. (2007). Citation analysis. *Annual Review of Information Science and Technology*, *41*, 609–641.

Paternoster, R., Brame, R., Mazerolle, P., & Piquero, A. (1998). Using the correct statistical test for the equality of regression coefficients. *Criminology*, *36*, 859–866.

Polanyi, M. (1962). The republic of science: Its political and economic theory. *Minerva*, *1*, 54–73.

Pratt, T. C., & Cullen, F. T. (2000). The empirical status of Gottfredson and Hirschi's general theory of crime: A meta-analysis. *Criminology*, *38*, 931–964.

Salman, M., Ahmed, M. M., & Afzal, M. T. (2021). Assessment of author ranking indices based on multi-authorship. *Scientometrics*, *126*, 4153–4172.

Schreiber, M. (2009). A case study of the modified Hirsch index hm accounting for multiple coauthors. *Journal of the American Society for Information Science and Technology*, *60*, 1274–1282.

Scopus. (2020). *Scopus content coverage guide*. https://www.elsevier.com/__data/assets/pdf_file/0007/69451/Scopus_ContentCoverage_Guide_WEB.pdf

Seo, C., Kim, B., & Kruis, N. E. (2021). Variation across police response models for handling encounters with people with mental illnesses: A systematic review and meta-analysis. *Journal of Criminal Justice*, 72, Article 101752.

Siebers, R., & Holt, S. (2000). Accuracy of references in five leading medical journals. *The Lancet*, *356*, Article 1445.

Simkin, M. V., & Roychowdhury, V. P. (2003). Copied citations create renowned papers? *[Preprint]*. https://arxiv.org/ftp/cond-mat/papers/0305/0305150.pdf

Simoes, N., & Crespo, N. (2022). On the measurement of scientific leadership. *Journal of Information Science*, *48*(1), 131–138. https://doi.org/10.1177/0165551520950240

Small, H. G. (1982). Citation context analysis. In B. Dervin & M. Voigt (Eds.), *Progress in communication sciences* (Vol. 3, pp. 287–310). Ablex.

Smith, L. C. (1981). Citation analysis. *Library Trends*, *30*, 83–106.

Smith, R. (1997). Journal accused of manipulating impact factor. *British Medical Journal*, *314*, Article 461.

Sweetland, J. H. (1989). Errors in bibliographic citations: A continuing problem. *The Library Quarterly*, *59*(4), 291–304.

Tewksbury, R., & Mustaine, E. E. (2011). How many authors does it take to write an article? An assessment of criminology and criminal justice research article author composition. *Journal of Criminal Justice Education*, *22*, 12–23.

Thornley, C., Watkinson, A., Nicholas, D., Volentine, R., Jamali, H. R., Herman, E., Allard, S., Levine, K., & Tenopir, C. (2015). The role of trust and authority in the citation behaviour of researchers. *Information Research: An International Electronic Journal*, *20*(3), 1–17.

Vavrycuk, V. (2018). Fair ranking of researchers and research teams. *PLOS ONE*, *13*(4), Article e0195509.

Wyles, D. F. (2004). Citation errors in two journals of psychiatry: A retrospective analysis. *Behavioral and Social Sciences Librarian*, *22*, 27–51.

Yang, D., Li, X., Sun, X., & Wan, J. (2016). Detecting impact factor manipulation with data mining techniques. *Scientometrics*, *109*, 1989–2005.

Zhang, L., Rousseau, R., & Sivertsen, G. (2017). Science deserves to be judged by its contents, not by its wrapping: Revisiting Seglen's work on journal impact and research evaluation. *PloS ONE*, 12(3), e0174205.

Author Biographies

John L. Worrall is a professor of Criminology and Criminal Justice at the University of Texas at Dallas. He has published articles and book chapters on a variety of topics ranging from legal issues in policing to crime measurement. He is also the author of several books, including the popular *Crime Control in America: What Works?* He currently serves as Editor of the journal *Police Quarterly* and as Executive Director of the Academy of Criminal Justice Sciences.

Ellen G. Cohn, PhD, is an associate professor in the Department of Criminology and Criminal Justice and an affiliated faculty member in the Center for Women's and Gender Studies at Florida International University. As well as studying scholarly influence and prestige in criminology and criminal justice, she currently is engaged in international cross-disciplinary research examining the relationship between crime and immigration. She is the Secretary of the Board of Directors of Miami-Dade Crime Stoppers and a recipient of the U.S. Presidential Volunteer Service Award.