Regular article

Do subjective journal ratings represent whole journals or typical articles? Unweighted or weighted citation impact?

William H. Walters

Mary Alice & Tom O'Malley Library, Manhattan College, 4513 Manhattan College Parkway, Riverdale, NY 10471, USA

A R T I C L E   I N F O

A B S T R A C T

This study uses journal ratings in criminology and criminal justice, library and information science, public administration, and social work to investigate two research questions: (1) Are stated preference (subjective) journal ratings more closely related to size-dependent citation metrics (eigenfactor and total citations, which represent the impact of the journal as a whole) or to size-independent citation metrics (article influence and CiteScore, which represent the impact of a typical article)? (2) Are stated preference ratings more closely related to unweighted citation metrics (five-year impact factor and source normalized impact per publication, which do not account for the impact of each citing journal) or to weighted citation metrics (article influence and SCImago journal rank, which do)? Within the disciplines evaluated here, respondents' subjective ratings of journals are more closely related to size-independent metrics and weighted metrics. The relative strength of the relationship between subjective ratings and size-independent metrics is moderated by subject area and other factors, while the relative strength of the relationship between subjective ratings and weighted metrics is consistent across all four disciplines. These results are discussed with regard to popularity and prestige, which are sometimes associated with unweighted and weighted citation metrics, respectively.

1. Introduction

Journal rating metrics—indicators of journal impact, prestige, reputation, utility, or perceived quality—can be readily classified into two types (Tahai & Meyer, 1999).[1] Revealed preference metrics are those that represent actual behaviors such as publishing, indexing, and citing. The most common revealed preference metrics are citation metrics such as the *h* index, impact factor (IF), source normalized impact per paper (SNIP), eigenfactor (EF), article influence score (AI), and SCImago journal rank (SJR). In contrast, stated preference metrics—also known as subjective or reputational ratings—represent scholars' opinions or hypothetical behaviors (e.g., "Which of these journals are most important to your work?" "Which carry the most weight in tenure and promotion decisions?"). Stated preference metrics are generally based on surveys of authors or faculty. They are most likely to be found in the social sciences and humanities, where the relationship between citation impact and perceived quality or reputation is not always straightforward. Moreover, stated preference metrics may better represent the opinions of scholars outside the "publish or perish" community—managers, policymakers, teachers, and industrial researchers, for instance (Bollen & Van de Sompel, 2008; Gorraiz & Gumpenberger, 2010; Schlögl & Gorraiz, 2010).

E-mail address: william.walters@manhattan.edu

[1] *Metrics* and *indicators* are used interchangeably here. Although *journal rankings* is a more common phrase than *journal rating*, the ratings themselves—rather than the ordinal rankings that result from them—are of primary interest in this study.

This paper uses multiple journal ratings in four disciplines—criminology and criminal justice, library and information science, public administration, and social work—to investigate two research questions:

(1) Are stated preference journal ratings more closely related to size-dependent citation metrics (those that represent the impact of the journal as a whole) or to size-independent citation metrics (those that represent the impact of a typical article)?[2]

(2) Are stated preference journal ratings more closely related to unweighted citation metrics (those that do not account for the impact of each citing journal) or to weighted citation metrics (those that do account for the impact of each citing journal)?

The first question relies on an important distinction. Size-dependent (whole journal) metrics such as total citations, EF, and the h index represent the number of citations accruing to all the articles in the journal. All else equal, a journal that publishes more articles will gain more citations, a higher EF, and a higher h index. In contrast, size-independent (typical article) metrics such as AI, CiteScore, IF, SJR, and SNIP divide total impact by the number of articles published and are therefore not influenced by journal size.[3] For citation metrics, the distinction between size-dependent and size-independent indicators is clear. With stated preference ratings, however, the instructions to survey respondents seldom specify whether they ought to be evaluating entire journals or a typical article within each journal. Section 4 addresses this question—whether scholars (respondents) consider journal size when rating journals.

The second question is based on the distinction between unweighted metrics (which assign equal weight to each citation, regardless of the characteristics of the citing journal) and weighted metrics (which assign higher weights to citations that appear in more influential journals). *Influence* refers to citedness and, in the case of SCImago Journal Rank, network centrality. Although nearly 20 unweighted and weighted citation metrics are available from data download sites such as Journal Citation Reports (JCR), Eigenfactor, CWTS Journal Indicators, SCImago Journal & Country Rank, Scopus Journal Metrics, and Google Scholar Metrics, it is not obvious that either unweighted or weighted metrics are preferable as indicators of impact, prestige, reputation, or perceived quality. Section 5 presents one way of addressing this issue; it identifies the type of indicator, unweighted or weighted, that more closely coincides with the journal ratings assigned by scholars.

These research questions are important for at least two reasons. First, investigations such as this can help us understand the relationships between impact, reputation, prestige, and related constructs as they apply to journals. We can use established citation metrics as landmarks, comparing them with stated preference ratings in order to better understand what survey respondents mean when they rate journals. This kind of comparison is possible, however, only if we first address the questions presented here. Second, comparisons of multiple metrics can help us gauge the convergent validity of each one. Newer indicators such as SNIP and SJR are more likely to be accepted if we know they are correlated with other indicators of journal "quality," especially when those other indicators use dissimilar methods to arrive at similar results (Cohn & Farrington, 2011; Martin, 1996; So, 1998; Weisheit & Regoli, 1984).

2. Previous research

Although many studies have investigated the correlations among citation metrics, fewer have examined the relationships between citation metrics and stated preference ratings. Two findings from the pre-2000 literature are especially notable:

(1) Stated preference ratings sometimes represent each journal's influence within a particular field or subfield rather than its more general scholarly impact. For instance, He and Pao (1986) discovered that the journal ratings assigned by scholars in the field of veterinary medicine are *inversely* related to the journals' impact factors ($r = -0.20$). However, those same ratings are directly related to the number of times each journal has been cited within a set of 74 leading veterinary journals ($r = 0.74$). This suggests that veterinary medicine is a relatively insular field in which journals are evaluated largely in terms of their influence on practice.

(2) The relationships between citation metrics and stated preference ratings are not always linear. In economics, sociology, and political science, for example, the top journals are assigned consistently higher subjective ratings than their IFs would suggest (Christenson & Sigelman, 1985; Ellis & Durden, 1991).

---

[2] The phrase *typical article* is used to distinguish size-independent metrics from size-dependent (*whole journal*) metrics. It is not strictly correct, however, since nearly every journal's citation distribution has a strong positive skew. For most journals, the average impact per article is substantially higher than the median impact per article (Calver & Bradley, 2009; Crookes et al., 2010; Seglen, 1997).

[3] The distinction between size-dependent and size-independent metrics has important implications for their use (Nisonger, 2004; Walters, 2016a, 2016b, 2016c). A librarian evaluating the cost effectiveness of various journals is likely to be interested in size-dependent metrics. In contrast, an author deciding where to send his or her paper may be more interested in size-independent metrics.