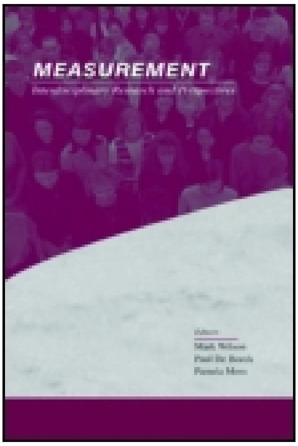

Measurement: Interdisciplinary Research and Perspectives

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hmes20

Metrics of Scholarly Impact

John T. Cacioppo [a] & Stephanie Cacioppo [a]

[a] Center for Cognitive and Social Neuroscience, University of Chicago

Version of record first published: 24 Sep 2012.

To cite this article: John T. Cacioppo & Stephanie Cacioppo (2012): Metrics of Scholarly Impact, Measurement: Interdisciplinary Research and Perspectives, 10:3, 154-156

To link to this article: http://dx.doi.org/10.1080/15366367.2012.720194

Measurement, 10: 154–156, 2012

ISSN: 1536-6367 print / 1536-6359 online
DOI: 10.1080/15366367.2012.720194

Psychology Press
Taylor & Francis Group

Metrics of Scholarly Impact

John T. Cacioppo and Stephanie Cacioppo
Center for Cognitive and Social Neuroscience
University of Chicago

Assessments of scientific contribution are important for a variety of reasons. For individuals early in their career, metrics of scientific work can provide valuable feedback about where they stand and the progress they have made. For departmental faculty seeking to hire a new member, such metrics can simplify the task of wading through hundreds of applications to identify a subset of applicants whom they might wish to interview. For departmental chairs, these metrics may influence annual raises and the allocation of limited resources. For university administrators, these metrics help identify faculty who warrant promotion and tenure, and departments that warrant concern. For scientific societies, these metrics influence the selection of award recipients across the course of careers. For public and private funding agencies, assessments of science help identify areas of progress and vitality that may warrant support as well as a means of documenting performance, ensuring accountability, and evaluating the return on their research investment. Measures of science can also be used for a variety of other purposes, such as to identify the structure of science, impact of academic journals, influential fields of research, and factors that may contribute to the likelihood of discoveries.

Ruscio and colleagues (Ruscio, Seaman, D'Oriano, Stremlo, & Mahalchik, this issue) provide a thoughtful empirical analysis of 22 different measures of individual scholarly impact. The simplest metric is number of publications, which Simonton (1997) found to be a reasonable predictor of career trajectories. There are numerous variants on this metric, such as the number of peer-reviewed articles, the number of first-authored publications, the number of articles published in a specific premier journal in a field, the number of articles without more-senior collaborators, and the number of publications across years (publication trajectories). Although the assessment of the scholarly contribution of a candidate for tenure is unlikely to be favorable if there are no or very few scientific publications, the number of publications is a far-from-perfect index of scientific contribution. Moreover, the incentives created by a metric such as number of publications may not promote the quality, innovativeness, programmatic nature, and cumulative impact of the research. Ruscio et al. (this issue), therefore, used five criteria to evaluate metrics of science: (1) ease of understanding, (2) accuracy of calculation, (3) effects of incentives, (4) influence of extreme scores, and (5) validity. Given the computational aids available today, the first 2 metrics are perhaps the least important and the third and fifth are most important.

Most of the indices examined by Ruscio et al. (this issue) are based on citation data, which may provide a better incentive structure for scientific contribution than simple publication counts.

Correspondence should be addressed to John T. Cacioppo, Center for Cognitive and Social Neuroscience, University of Chicago, 5848 S. University Avenue, Chicago, IL 60637. E-mail: jcaciopp@uchicago.edu

By indexing reference lists of articles published in journals, the Institute for Scientific Information (ISI) has made it simple to determine the number of times any given article, body of work by an author, or articles appearing in a particular journal were cited in published journal articles. Although errors can occur, indexing citations is generally objective and reliable, and it can be calculated excluding self-citations to provide a metric of the extent to which a work or body of work has influenced others. One of the interesting results of the empirical analyses performed by Ruscio et al. is that self-citations have nominal impact on the outcomes, a finding that increases the confidence in the validity and simplifies the calculation of these metrics.

The *source* of publication and citation data is an important consideration. Limitations in the ISI dataset (Web of Knowledge), such as the typical absence of books and chapters, may be filled by other citation databases including Scopus and Google Scholar. Scopus provides better coverage than Web of Knowledge of conferences, but the coverage of publications prior to 1992 is poor. Google Scholar has good coverage of conferences, books, book chapters, and most journals, but the coverage of journals prior to 1990 is somewhat limited and work that is not available through archival publication channels (gray literature) is included. This broader coverage usually leads to higher citation counts using Google Scholar than using Web of Knowledge for searches in the social sciences, humanities, and engineering. At this juncture, Google Scholar's coverage of some fields of science is less complete than ISI and some publishers do not permit open access to their journals for one year, which delays their availability through public search engines such as Google Scholar. Ruscio and colleagues used PsycINFO, restricted to peer-reviewed journal articles, as a source of citation data, a decision that may bias their findings. It would be helpful to know how their results might differ when Google Scholar or Scopus are used as a source of data.

Perhaps the best metric to emerge in Ruscio et al.'s (this issue) analysis is the *h* index (Hirsch, 2005). This metric provides a single number that balances the number of publications and the number of citations per publication. The *h* index has limitations, of course. Among them are that the *h* index tends to increase with years as a scientist; gratuitous authorship can contribute to inflated scores; the *h* index does not take into consideration the number or the role of the authors; different citation databases provide different *h* indexes as a result of differences in coverage; the *h* index is bounded by total number of publications; the *h* index does not consider the context of the citations (e.g., negative findings or retracted work); and individuals with the same *h* index may nevertheless differ dramatically in total citations or in number of publications. The *g* index, proposed by Leo Egghe (2006), attempts to address the last concern by giving more weight to highly cited articles. Given that a set of articles are ordered in terms of decreasing citation counts, the *g* index is the largest number such that the top-*g* articles together total at least g^2 citations.

The availability of quantitative data makes metric-based decision making simpler than ever. Although some metrics performed better than others, which metric is preferred should depend on the question that is asked. For instance, Ruscio et al. (this issue) found that the *h* and *g* indices perform quite similarly in their analyses, so the *h* index may be preferred based on ease of understanding and calculation, whereas, the *g* index might be preferred when evaluating the contributions of scientists with similar *h* indices. Moreover, citation data may not be as helpful when evaluating candidates for the position of beginning assistant professor as when evaluating candidates for promotion to full professor. As in the case of operationalizing any theoretical variable, using multiple operationalizations may be advantageous. When multiple metrics of science converge on the same result, confidence in the result is increased. When these metrics provide different outcomes, discrepancies across the indices can be used to draw more-informed interpretations.

Citation data have noteworthy limitations as well. As objective and replicable as the results might appear, the limitations of citation data include that articles appearing in early issues of new journals may not be indexed; books, chapters, and published conference proceedings are seldom or irregularly indexed; authors whose work becomes so well accepted, they no longer are cited for the discovery; errors in references ranging from the misspelling of author names to the misspecification of the page numbers or author initials contribute to underestimates of citation impact; authorship order is ignored; tutorials of methodological or statistical techniques may garner large citation counts compared to the groundbreaking work that led to the development of the technique; articles in faddish or insulated areas may have a sizable short-term influence but little if any long-term influence; insulated groups of investigators can collude to cite each other's work as a means of inflating citation counts; fields differ in terms of conventions concerning references to existing literatures, half-lives for published work, and lags between submission of manuscripts; and publication of articles vary across journals, fields, and disciplines. Ruscio et al.'s (this issue) evidence for the validity of modern citation-based metrics is encouraging but additional work is needed to develop even better objective metrics.

Finally, metric-based decision making can also have the unintended effect of promoting scientific work that yields higher values on the selected metric rather than more meaningful, innovative, and/or cumulative scientific work. For this reason, quantitative metrics should not substitute for reading a scholar's work and speaking with the scholar about that work. Doing the latter can help determine whether a person's publications are advancing or cluttering a field and can help disambiguate whether a person's contribution to an influential series of publications is limited or transformative. Therefore, a faculty's reading of a body of scientific work, scientific presentations and discussions with an author or authors, and views expressed by external experts are likely to continue to serve an important role in the evaluation of body of scientific work. The need for such approaches is evident, for instance, when measures of science are used in macro analyses, such as to determine influential fields of research that warrant funding or support, or to provide a means of evaluating the return on research investments.

REFERENCES

Egghe, L. (2006). Theory and practice of the g-index. *Scientometrics*, *69*, 131–152.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, *102*(46), 16569–16572.

Simonton, D. K. (1997). Creative productivity: A predictive and explanatory model of career trajectories and landmarks. *Psychological Review*, *104*, 66–89.