aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Perspectives on Psychological Science
2016, Vol. 11(6) 905–908

DOI: 10.1177/1745691616664436
pps.sagepub.com
SAGE

Taking Advantage of Citation Measures of Scholarly Impact: Hip Hip *h* Index!

John Ruscio
The College of New Jersey

Abstract

Professional decisions about hiring, tenure, promotion, funding, and honors are informed by assessments of scholarly impact. As a measure of influence, citations are produced by experts but accessible to nonexperts. The *h* index is the largest number *h* such that an individual has published at least *h* works cited at least *h* times apiece. This is easy to understand and calculate, as or more reliable and valid than alternative citation measures, and highly robust to missing or messy data. Striving for a large *h* index requires both productivity and influence, which provides healthy incentives for researchers striving for eminence through scientific impact. A number of factors that can influence *h* are discussed to promote the mindful use of what might otherwise be an ambiguous or misleading measure. The *h* index adds a transparent, objective component to assessments of scholarly impact, and even academic eminence, that merits at least two cheers.

Keywords

scholarly impact, fame, *h* index, citations

High-stakes decisions such as hiring, tenure, promotion, and the awarding of research funding or professional honors are often made on the basis of scientific impact. Citation measures can play a valuable role in assisting those charged with making these decisions. Particularly for individuals outside a candidate's area of expertise, which includes most administrators and many professional colleagues, a well-chosen metric can be a useful complement to a CV and external review letters.

As a data source, citations offer several advantages relative to traditional measures. Counting publications rewards being prolific even if no real influence is achieved. Ratings of scholarly outlets are not specific to an individual's work. Even at leading journals, a small fraction of the articles exerts most of the influence. Citations track the influence of articles in a more fine-grained manner than coarse measures like journal impact factors. Citations draw from the specialized knowledge of professional colleagues, yet no such expertise is required to use citation measures.

What the *h* Index Brings to the Table

One highly cited publication is impressive but less so than the sustained production of many influential works. Unlike the total citation count, many citation measures are robust to outliers in the distribution of citation counts. The most popular index is easy to calculate and understand: The *h* index is the largest number *h* such that an individual has published at least *h* works cited at least *h* times apiece (Hirsch, 2005). Geometrically, *h* is the length of the largest square that fits within an array of publications by citations. Merely being prolific will not lead to a large *h* index, nor will a perfectionistic focus on a small number of scholarly gems. Growing the *h* square is two-dimensional and requires the sustained production of influential work.

Many alternative measures have been proposed, but there are a number of reasons to prefer the *h* index (Ruscio, Seaman, D'Oriano, Stremlo, & Mahalchik, 2012). Citation measures tend to be very highly correlated, and *h* is at least as reliable and valid as its rivals. Moreover, it is the easiest index to calculate and understand as well as the index most robust to outliers or missing data. Due to space constraints, only *h* will be discussed further (see

Corresponding Author:
John Ruscio, Department of Psychology, The College of New Jersey, P. O. Box 7718, Ewing, NJ 08628
E-mail: ruscio@tcnj.edu

Ruscio et al., 2012, for pros and cons of many citation measures).

Decision makers are paying increasing attention to the *h* index, which reinforces desirable goals among scientists who aspire to eminence by making important scientific contributions. The *h* index measures more than productivity because scholars must also earn recognition through citations. In addition, *h* is a transparent, reproducible, and objective measure of scholarly impact. This has the potential to reduce many kinds of bias that can influence judgments and decisions. It also affords the opportunity for systematic, quantitative evaluations, or comparisons. Unlike the more subjective reading of a CV, for example, one can explicitly select an appropriate reference group and sampling technique to obtain comparison data to help interpret the *h* index.

Consider the case of social psychologist Flurpple applying for promotion to full professor at an elite research university. Flurpple creates a Google Scholar profile that calculates an *h* index of 33. Is this compelling evidence that supports the application for promotion? To illustrate how Flurpple could obtain pertinent data, I found a list of top social psychology programs and obtained the *h* index for all 104 tenure-line psychologists who were affiliated with these programs and had Google Scholar profiles.[1] Doing so took less than 2 hours (see Harzing, 2013, on the efficient search for and appropriate uses of citation data). Figure 1 shows the distributions of *h* for 17 assistant professors, 20 associate professors, and 67 full professors. Flurpple's score of 33 exceeds all but one of the associate professors. Though it would fall within the bottom quartile of full professors, that would be expected for a newly promoted individual. Actual candidates for hiring, tenure, promotion, or other professional recognition or awards could be asked to furnish comparison data tailored to suit the circumstances. This illustration demonstrates the feasibility of the approach.

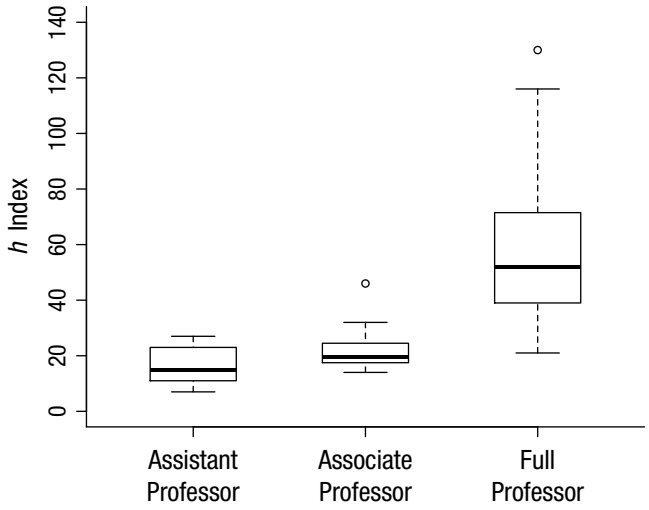

Fig. 1. Boxplots showing score distributions on the *h* index retrieved from Google Scholar profiles for 17 assistant professors, 20 associate professors, and 67 full professors affiliated with highly regarded social psychology programs. An "o" represents an outlier.

Mindful Use of the *h* Index

The *h* index holds great promise as one component of an assessment of scholarly impact. Its simplicity is a virtue, but there are no magic numbers that constitute generally applicable thresholds for good, excellent, or outstanding scores. A number of factors that can influence *h* or its interpretation must be considered carefully.

First, an individual's *h* index depends on what database is used to retrieve citations, when, and how effectively the search is performed. Calculating *h* is a transparent and reproducible process, but the citation data themselves are not necessarily free of error or bias (see Eagly & Miller, 2016, this issue, on potential gender bias). Significant work has been done to catalog citations electronically. Google Scholar has evolved dramatically and is arguably the database of choice for retrieving citations in any academic discipline (Harzing & Alakangas, 2016). Advantages include free access, ease of use, speed, and comprehensive coverage. Coverage is especially important in disciplines such as the social sciences, in which scholarly work often appears in journals, books, chapters, conference presentations, or other outlets not indexed elsewhere. Harzing and Alakangas (2016) showed that for social scientists, Web of Science or Scopus might provide as few as 20% to 30% of the citations in Google Scholar.[2] The chief disadvantage of Google Scholar is that its automated processing can misidentify authors and provide messy data. Fortunately, one strength of the *h* index is its robustness to missing or messy data (Harzing & Alakangas, 2016; Ruscio et al., 2012).

Second, the *h* index only takes integer values, so tied scores will be common. For example, only 60 unique values of *h* occurred in the sample of 104 professors described earlier. Ties can be a nuisance, and many alternatives to the *h* index—including traditional, subjective reviews of the scholarly record—produce few or no ties. This apparent precision, however, is extremely likely to be illusory and swamped by various sources of measurement error. It is crucial not to make too much of small differences on any measure, and we must accept that important decisions will sometimes be made within the margin of error.

Third, there is concern about awarding credit for self-citations. Perhaps because self-citations tend to be strongly correlated with total citations, it may not be surprising that Ruscio et al. (2012) found that the reliability of citation indices failed to improve when self-citations (operationalized in any of five ways) were removed.

Given the difficulty of identifying self-citations, there appears to be no compelling reason to do so. A related concern is that self-citation could be done strategically to flesh out the empty corner of an *h* square. This would require impressive foresight about which scholarly works will occupy that position in the long run. The publications by citations array evolves in unpredictable ways.

Fourth, there is a similar concern about the merits of awarding full credit for citations of works with shared authorship. Ruscio et al. (2012) found that the reliability of citation indices failed to improve after adjusting for multiple authors (using any of four different algorithms), once again suggesting no compelling reason to do so. Imposing a one-size-fits-all penalty on citation credit for shared authorship fails to acknowledge the many circumstances surrounding collaborative publication and could easily introduce more problems than it solves. It seems more prudent to consult the CV to identify cases in which an individual is seldom the lead author or often has many coauthors.

Fifth, because citations accumulate over time, individuals at later career stages would be expected to have larger *h* indices. If necessary for a particular purpose, one can control for career stage by dividing *h* by the number of years since the first publication. This yields what Hirsch (2005) called the *m* quotient. Unlike *h*, it can decline over time if scholarly impact wanes. For example, if the *h* square does not quadruple in size over a doubling of career length (e.g., from $h = 10$, or 100 citations in the square after 10 years, to $h = 20$, or 400 citations in the square after 20 years), the *m* quotient will decline. Thus, to compare individuals at different career stages, the *m* quotient demands sustained productivity and influence.

Sixth, perhaps the most vexing problem with *h* involves the differences across scientific disciplines or even subdisciplines or specialty areas. These differences arise for many reasons, such as the number of people working in an area, the extent to which that work reaches beyond specialists, the size of research teams, and the typical number of references for publications in a certain field of study. With so many sources of variance, comparing individuals who work in different areas can be unfair. Of course, the same challenges plague any attempt to assess scholarly impact, not just a citation measure. It should be easier to properly contextualize a comparatively simple, objective measure like the *h* index than measures that are more complex or subjective. I am not optimistic that global norms for *h* can be developed in the foreseeable future, but the technique illustrated above for Flurpple's promotion application shows one approach that is feasible for high-stakes decisions. A candidate can be asked to produce and document a sample of comparison data to help interpret his or her *h* index. In time, it might become feasible to automate much of what I did manually.[3]

Conclusions

The *h* index does not incorporate all pertinent information and using it will not solve all decision-making dilemmas. That is too much to ask of any measure. However, the *h* index is transparent, objective, and based on a wealth of information that usefully augments what can be gleaned from a strictly subjective review of a scholarly record. It is both possible and, increasingly, feasible to obtain an appropriate sample of comparison data to help interpret the *h* index. For these reasons and in recognition of its other desirable characteristics, Hirsch's (2005) *h* index deserves at least two cheers.

Familiarity with and use of the *h* index is accelerating rapidly. Knowing that your *h* index probably will be evaluated one or more times during your career may have implications for how to go about your work. However, unless you have been following an ill-advised path, the advice should not be troubling or even surprising: Aim for the sustained production of influential work. We are leaving behind the days of mindlessly counting the number of publications on a CV. In the early stages of an academic career it may remain tempting to publish as much as possible, but as your career unfolds you will have the luxury of caring as much about impact as productivity. The *h* index helps to assess the success with which both are attained and its use might help to steer scholars toward the more laudable goal of scientific impact.

Can the *h* index help to answer the question, "Am I famous yet?" Perhaps, if you want to operationalize fame as scholarly impact. It is not difficult to calculate your own *h* index using Google Scholar, nor to obtain comparison data to help make sense of this otherwise highly ambiguous number. Who is your reference group? At what percentile do you consider "fame" to have been achieved? Ultimately, whether this is a worthwhile exercise depends on how you will manage the inevitable trade-offs of your professional and personal life. As Roediger (2016, this issue) notes, fame in psychology comes in many varieties and can be fleeting. What are your research, teaching, and service goals? What kinds of relationships do you want to have with family, friends, and community? What steps will you take to bring these goals to fruition? When it becomes necessary to prioritize, what sacrifices will you make in one area to support another? The *h* index can quantify scholarly impact and help to identify academic eminence, but it cannot tell you how heavy an emphasis you should place on its pursuit.

Declaration of Conflicting Interests

The author declared no conflicts of interest with respect to the authorship or the publication of this article.

Notes

1. This list of social psychology programs was obtained from the Social Psychology Network at https://www.socialpsychology.org/gsocial.htm but originates with Gourman (1997). I wanted to include at least 100 professors, and I reached that target by the 20th of the 32 programs on this list. My use of this list for illustrative purposes should not be construed as an endorsement. Gourman used an undisclosed methodology (see the critique by Bedeian, 2002), and this list is, in any event, now quite dated. An actual candidate for promotion could do much better by consulting a list of programs considered by colleagues at his or her own institution to be appropriate for comparative purposes. Also note that I limited my search to individuals with Google Scholar profiles as a means of quality control. For those who have not created a profile, particularly for those with relatively common names, it is more difficult to determine which Google Scholar search results should be included in the calculation of the *h* index.
2. Ruscio and Prajapati (2013) found that PsycINFO and Web of Science yielded very similar *h* indices for psychological scientists, suggesting that PsycINFO misses a comparably large proportion of citations.
3. If zillow.com can figure out which recently sold homes are comparable enough to ours (e.g., based on location, age, and size) to produce decent estimates of our home values, then Google Scholar (or another service) should be able to figure out which scholars are comparable enough to us (e.g., by virtue of working at similar institutions and publishing in similar disciplines) to produce decent samples of citation measures with which to evaluate our scholarly impact.

References

Bedeian, A. G. (2002). *Caveat emptor*: The Gourman Report. *The Industrial-Organizational Psychologist, 73*. Retrieved from http://www.siop.org/tip/backissues/tipjan02/07bedeian.aspx

Eagly, A. H., & Miller, D. (2016). Scientific eminence: Where are the women? *Perspectives on Psychological Science*, *11*, 899–904.

Gourman, J. (1997). *The Gourman Report* (8th ed.). New York, NY: Random House.

Harzing, A. W. (2013). *The publish or perish book*. Retrieved from http://www.harzing.com/popbook/index.htm

Harzing, A. W., & Alakangas, S. (2016). Google Scholar, Scopus, and the Web of Science: A longitudinal and cross-disciplinary comparison. *Scientometrics*, *106*, 787–804. doi:10.1007/s11192-015-1798-9

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences, USA*, *102*, 16569–16572. doi:10.1073/pnas.0507655102

Roediger, H. L. (2016). Varieties of fame in psychology. *Perspectives on Psychological Science*, *11*, 882–887.

Ruscio, J., & Prajapati, B. (2013). Citation-based indices of scholarly impact: Databases and norms. *APS Observer*, *26*(7). Retrieved from http://www.psychologicalscience.org/index.php/publications/observer/2013/september-13/citation-based-indices-of-scholarly-impact-databases-and-norms.html

Ruscio, J., Seaman, F., D'Oriano, C., Stremlo, E., & Mahalchik, K. (2012). Measuring scholarly impact using modern citation-based indices. *Measurement: Interdisciplinary Research and Perspectives*, *10*, 123–146. doi:10.1080/15366367.2012.711147