Exploring Impact Metrics Beyond Indices and Citations for SOTL and Applied Educational Research

Melissa S. Medina, EdD, a JoLaine R. Draugalis, PhD a,b

a University of Oklahoma College of Pharmacy, Oklahoma City, Oklahoma
b Editorial Board Member, American Journal of Pharmaceutical Education, Arlington, Virginia

Submitted June 13, 2018; accepted July 20, 2018; published April 2019.

Keywords: research, peer review, academic promotion, metrics, publications

One of the responsibilities of academic pharmacy faculty is to demonstrate excellence within the tripartite mission of teaching, research and service. Faculty members document their activities and achievements at least annually and periodically for promotion and/or tenure reviews. Demonstrating excellence in research may appear to be the easiest of the tripartite mission since faculty members can quantify their scholarship. They can present total number of research publications in peer reviewed journals, number of citations per paper, number of book chapters, number of grants as principal investigator, and total grant dollar amount. But in the words of Albert Einstein, “not everything that counts can be counted, and not everything that can be counted counts.” While these metrics offer a way to “quantify” research accomplishments, administrators and peer reviewers are also concerned about the “quality” of the journals faculty publish in and the “quality” of the papers that cite the faculty member’s research.

A faculty member can evaluate his/her research impact and visibility using author metrics or article metrics. Author metrics can be found in Scopus (by Elsevier), Web of Science, Google Scholar Citations, ImpactStory, ResearchGate, Academia.edu, and Plum Analytics; whereas article metrics are offered from these aforementioned sites as well as Mendeley, PLOS, and Altmetric. There are six main metrics used to evaluate quality: impact factor (IF), h-index, i10-Index, g-index, Eigenfactor, and Article Influence Score (AIS), but others exist. Impact factor is one measure of journal quality. Developed by Eugene Garfield in 1995 as a tool for librarians to manage library collections, it is now used to rank and evaluate journals based on the number of citations for an average article for a given journal and is also used as a tool for evaluating the performance of individual researchers. Impact factor is calculated by dividing the number of times articles (in indexed journals) were cited by the number of articles that are cited over a two-year period. Although impact factor is not a direct measure of author impact, it is considered a relevant metric because it offers insight into the quality of the journal(s) in which an author publishes. There is an ongoing debate about the reliability of impact factor to measure journal quality since it can suffer from miscalculation and discipline bias.

For example, pharmacy, health education, and nursing are regarded as low profile disciplines in impact factor because there are only a limited number of Institute for Scientific Information (ISI) ranked journals in each discipline. Another limitation of the impact factor is that the number of citations calculation includes self-citations by the author, which skews the data.

In addition to the impact factor, there are other metrics to evaluate journal and article quality such as (but not limited to) the h-index, i10-Index, g-index, Eigenfactor, and the Article Influence Score (AIS). A brief explanation of these five metrics is included to better orient the reader to existing measures used in the Academy to judge quality; however, more detailed explanations of each metric are available elsewhere.

The h-index measures productivity and impact of an individual author’s published work (versus journal ranking) and is calculated by ranking an author’s articles in decreasing order of citations, with the h-index being the highest number of papers that received h or more citations, but it is limited by the metric’s inability to capture author order or changes over time. The i10-Index is the number of publications with at least 10 citations and although it is simple to calculate, it is only used by Google Scholar.

The g-index measures an individual researcher’s quality outcomes; the higher the number of citations received for an article, the higher the g-index, therefore it is sensitive to highly cited articles. Eigenfactor scores are intended to measure how likely a journal is to be used by the average researcher; it is similar
to impact factor but takes into account self-citation.\textsuperscript{1} The AIS is similar to a five-year impact factor as it is the average influence of a journal’s articles over the first five years after publication, but can’t be calculated for recently published articles.\textsuperscript{1}

Overall, faculty members use these metrics to demonstrate their publication quantity and quality (although limited) for annual review or promotion and/or tenure. But one metric is missing for faculty who publish scholarship in teaching and learning (SOTL) or applied educational research. This metric could be called “article utility.” This “article utility” metric would capture how “useful” an article is to other faculty or administrators who are designing or assessing programs, courses, or lectures, but are not publishing an article about how they “used” the initial article. For example, an individual faculty member (A) may have published an article that is used by another faculty member (B) to design his/her course at his/her college or school of pharmacy, but faculty member B does not publish the results of his/her course outcomes in a journal. As a result, the quality of faculty member A’s scholarship is not captured or quantifiable because there is no public citation of the work. One example of a SOTL course/program design and assessment article that illustrates this limitation is by Medina and Draugalis who published a teaching philosophy article about best practices for writing and evaluating teaching philosophies. The article included a rubric.\textsuperscript{8} The authors received anecdotal evidence from outside department chairs about how the article is required reading at their institution for faculty to use the article to write or revise their philosophy for annual review, promotion and/or tenure, and awards dossiers. Residency program directors would also tell the authors that they use the rubric to evaluate their residents’ teaching philosophies. In these two examples, the products that resulted from the teaching philosophy article are not publishable in the literature (a revised or graded teaching philosophy) and therefore, are not citable, which prevents the products from being included in the traditional quality measures of an article described earlier (eg, impact factor, h-index, g-index). However, it seems apparent that the article’s advice or assessment tool affected others’ work products. When peers use the content of an article to inform how they individually or institutionally improve their work products without publishing in a journal that would allow them to cite the work they used, it makes it difficult for the authors to demonstrate the quality or impact of their work. For example, if a department chair or promotion and/or tenure committee were evaluating the “impact” of the teaching philosophy article, they would note that the article has been cited six times on Google Scholar but read 552 times according to ResearchGate.\textsuperscript{9} ResearchGate is a free online networking site for researchers that allows researchers to track the number of “reads” for a given article.\textsuperscript{9} This is just one, but not uncommon, example of faculty who publish in SOTL.

A second example of this missing metric is for curricular design and process articles; articles about how to conduct a program and the resources and plans that are needed. Medina and Castleberry published a SOTL article about strategies for proctoring online versus paper examinations.\textsuperscript{10} The paper outlined how to revise existing test questions, use a seating chart, manage student belongings, write academic misconduct statements, and monitor students during the test. The paper also provided suggested policies or guidelines programs could use to deliver electronic examinations that would uphold academic integrity, such as using a test password, test start and stop time, seating chart, signed test scratch paper, and a device maintenance checklist for students to use prior to entering the examination room. A review of the article quality indicators statistics revealed that the paper has not been cited in Web of Science, but has been read 23 times on ResearchGate. The impact of this paper may not be seen in new research publications about the topic, but it may influence a school’s syllabus template, academic misconduct statement, and/or examination proctoring training. The work may be a “fleeting reference,” that which is cited in local residency or faculty development manuals or teaching tips newsletters, or on college or university office of instruction websites, which are citations that are difficult to know exist, let alone capture. However, the “article utility” still exists. In the example of the examination proctoring article, procedural guidance in this area is important because it is reported that cheating in the health sciences is prevalent and often remains undetected by faculty.\textsuperscript{11-13} Programs and/or faculty who use this article would not publish their syllabi in the scholarly literature but that does not diminish the influence the article has had on all of the program’s stakeholders, including students.

These two examples highlight the limitations of existing quality research metrics in capturing research or SOTL use. Therefore, additional alternative metrics, such as “article utility” are needed for faculty when they are summarizing their scholarly efforts. One option for faculty who have published SOTL articles designed to influence program or course design, delivery or assessment or program procedures is to create a profile of their research in ResearchGate to track the number of “reads” for their articles.\textsuperscript{9} This metric shows how often the work was accessed in ResearchGate, including each time a person [other than the article author(s)] downloads the article or views the article’s summary or figures.\textsuperscript{9} A limitation of this alternative metric is that the number includes the number of downloads, but does not indicate if the person who downloaded
the article actually read the article. Another limitation is that it only captures those who have read the article using ResearchGate, versus the actual print or online version from the journal itself or the reader’s college or school library, so the number of “reads” may be underreported. This metric also does not capture how the article influenced the reader’s work on the topic after he/she read the article. A final limitation is that the number of reads is influenced by the availability of the author’s full-text articles, but putting a copy of an article on ResearchGate may violate copyright policies for certain journals. Similar to the number of reads in ResearchGate is the number of recommendations, which captures the number of times a ResearchGate reader encouraged others in her/her network to read the article. This is an additional data point, but less insightful than the number of reads. Also available on ResearchGate is an RG score, which is a metric that measures scientific reputation based on how a faculty member’s work (published research and contributions to ResearchGate) is received by peers. It is an overall research score that includes percentages related to the faculty member’s publications, followers, and questions asked and answered on ResearchGate. The RG score can be broken down into a percentile rank compared to other ResearchGate members. The limitation of the score is that it is not a score for a single article, but all of the author’s research, which could be skewed by one or a limited number of high performing articles. RG scores only consider contributions to ResearchGate by those who are members; it does not include judgement of one’s reputation from outside members, which presents another limitation. The RG score is also influenced by the availability of the author’s full-text articles, which again, may make authors in breach of publisher policy.

Another alternative metric that faculty could quantify for article use is the number of times the SOTL or research article is presented locally, nationally, or internationally, whether these presentations were invited or accepted, and the presentation setting such as a workshop, retreat, symposium, or conference. This grouping of metrics could allow evaluators to better understand how the faculty member’s article is being “used” by the Academy or its quality or value-add. For example, using the examination proctoring example described previously that had no citations, one of the authors has frequently presented on this topic using the article and could quantify the presentations as one national invited presentation, two national invited faculty workshops, 10 invited university training workshop presentations, 15 college training workshops; total equals 28 presentations.

Given the importance of interprofessional practice (IPE) in the health professions, another metric that faculty could quantify is the applicability or utility of a SOTL article for an interprofessional audience. In the examination proctoring example, the article could be easily used in any health science or general education audience since other professions offer examinations to students and need to maintain academic integrity in the testing environment. The number quantified could be a sum of the presentations that could use the article or it could be an IPE yes or IPE no indication.

Overall, a faculty member’s work may not be cited in research publications because the article may be focused on best practices on how to design, deliver, teach, or assess an area. Others who use the article to improve their program or lecture may be scholarly teachers who are not looking to publish their results. The Academy needs to find alternative metrics, such as an “article utility” measure, that faculty members can use to demonstrate the impact of their SOTL on programs, curricula, courses, and students. While three suggestions have been offered, continued dialog needs to occur.

REFERENCES