MAKING AN IMPACT

The Impact Factor’s Intent, Benefits, Limitations, and Competition

Presented by Erin Owens
Associate Professor, Newton Gresham Library, SHSU
Access Services Coordinator & Scholarly Communications Librarian

5 Oct 2018
What is *impact*?

- The Journal Impact Factor (IF)
  - Definition and Intent
  - What it Can (and Can’t) Do
  - Problems
- Many Metrics for Many Purposes
- Limitations of Metrics
- Telling the Story of Your Research
- Wrap-Up and Q&A
When I say *impact* in the context of scholarship, what does that mean to you? [Activity during session: Discussion]

Ultimately, the definition of the term undergoes variations, sometimes subtle and sometimes stark, based on context. A public health researcher and a literature researcher likely have very different ideas of what “impactful scholarship” looks like.

To provide a foundation for our discussion, I will borrow a definition from the book *Meaningful Metrics* by Roemer and Borchardt (2015); they define impact as including two principles: “*effect*, in the sense of a perceptible shift, change, or influence” and “*force*, in the sense of the strength or degree of this effect,” thereby yielding “a two-part determination of where a work can be said to have an effect and to what extent the force of this effect can be quantified and benchmarked” (4).
Impact Factor (IF) is a citation-based ratio—it considers the number of citations from other considered journals that occurred in the current year to articles published in the previous two years, divided by the total number of “substantive articles and reviews” published in the same two years.

IF was devised by this man, Eugene Garfield, a PhD Linguist, the creator of the Science Citation Index, the founder of ISI (the Institute for Scientific Information), and one of the founders of the field of bibliometrics and scientometrics. Journal Impact Factors have been consistently published via the Journal Citation Index since 1975.

[Activity during live presentation: Do any elements of this ratio raise questions for you, or do you find any aspects of it problematic? Ask attendees for thoughts; discuss. As needed, draw on points below, which are narrated in detail for those reading the slides after the presentation.]

Let’s start with “number of citations” in the ratio’s numerator. This does not comprehensively include all possible citations to a work, but rather, it is limited to the scope of citations from other scholarly journals that are also indexed in Web of Science. This is probably a reasonable perspective for those doing the work, but consumers of this metric need to understand that the number of citations counted is necessarily a sampling.
What about “substantive articles and reviews,” a quote I have taken from Garfield? The term is problematic, because it is not clearly defined in a way that others can reliably replicate; what articles, editorials, reviews, or other pieces in a journal issue are included in or excluded from the denominator in this ratio?

The two-year timeframe raises other questions. Again, from the purpose of those doing the work to compile this information, it is necessary to place some sort of boundaries on the chronological scope. The initial focus was on the hard sciences, and two years was selected as a reasonable period that would allow citing research to be completed, reviewed, and published, while still emphasizing currency. Garfield has stated that, “The impact factor could just as easily be based on the previous year's articles alone, which would give even greater weight to rapidly changing fields. An impact factor could also take into account longer periods of citations and sources, but then the measure would be less current.” However, even acknowledging that some time limit is needed, the appropriateness of the two-year window will likely vary between disciplines. Additionally, slow review processes at a journal may have a decreasing effect on the IF, because some citations won't appear until after the 2-year window. I will note that there is also a 5-year Impact Factor provided in Journal Citation Reports, which may be more helpful in some disciplines but still may not appropriately reflect citation habits outside of the hard sciences.
It is important to realize that the Journal Impact Factor is a proprietary metric, a trademark owned by Clarivate Analytics (formerly owned by Thomson Reuters).

This number is not an inherent, objective fact—a profit-making entity ultimately determines what elements factor into the calculation, so it is not guaranteed to be transparently reproducible, and no other entity can independently calculate or publish this number.
Now that we understand what IF is, let’s take a moment to discuss its original intent and some of its extended uses.

What Garfield was really focused on was creating his Science Citation Index—a clear view of who cited whom in order to quickly identify all papers that cited or criticized another paper. The Impact Factor ratio was developed to help select highly cited journals for coverage in this index. They could have compared total citation counts, but this ratio was determined to be more fair to small, specialized, yet nevertheless important journals that should be considered. So that’s it—the whole original purpose was just to define the scope of the Science Citation Index, which we now know as the Web of Science database (WOS).

The metric also became a tool to help libraries determine how to allocate limited funds for journal subscriptions by selecting those that seemed to be garnering the most attention in a discipline. In the pre-Internet age, citation was the only obvious metric available to measure such attention.

And, of course, the metric helps researchers to compare the relative “citedness” of journals within their discipline.
What It Can Do

IF can compare journals in WOS...

Within a single field...

To see which receive more citations...

On average...

From other journals in WOS...

Within a 2-year (or 5-year) period.
It is very important to discuss the things that IF cannot do.

- IF cannot speak to all disciplines equally. Disciplines in the hard sciences tend to receive better coverage than the social sciences, and fields in the arts and humanities are not addressed at all. Thus, while it may be a stronger tool for judging journals in some fields, it may be weaker or entirely inadequate in other fields.

- IF cannot address new journals that have not yet been published for three years, or journals in niche specializations which may be critical to researchers in that area but which due to the small scope will never “rate” the minimum number of citations for WOS inclusion.

- IF is not an indication of the literal quality of a journal—it can only measure discussion of that journal’s work relative to others in the same field.

- IF cannot be compared across fields. Citation behaviors, and thus citation rates, vary too greatly between disciplines for their IF’s to be comparable across discipline lines.

- IF pertains to the journal level only and should not be treated as a proxy for evaluating a specific article, researcher, or institution.

- Some disciplines or researchers produce work intended to influence practice, policy, or the public, and impact in those areas may not be reflected in scholarly citations, so IF
may be less informative in those fields.

It is also worth noting that some disciplines may be combined in JCR in ways that do not logically “fit” with the way research is actually conducted or with the audience for those journals, which may skew ratios and rankings for the journals in a specific research area. As just one example, Thomson compiles journals from both Management Information Science (MIS) and Library Information Science (LIS) disciplines into a single category called *Information Science & Library Science*; however, these fields represent two distinctly different groups of researchers. It unfairly warps perspective to judge the 1.3 Impact Factor of this category’s 24th-ranked *College & Research Libraries*—published by the Association of College and Research Libraries, a division of the American Library Association—against the 5.4 Impact Factor of the 1st-ranked journal in the category—*MIS Quarterly*, published by the University of Minnesota School of Management—when there is a distinct separation and little overlap between the populations contributing to and consuming these publications.
We’ve touched on a few of IF’s limitations already – The citations considered are only from other journals within the Web of Science, and what exact items in an issue are included in the calculation is not necessarily transparent or reproducible. The emphasis on the recentness of citations may be less effective for some disciplines (e.g., sciences versus humanities), and the classification of journals into disciplines may not always be ideal.

I’d like to address a few other concerns as well.

First, IF gauges discussion of the work in a journal, but discussion is not always positive, so treating the IF as a proxy for quality or prestige (as often happens when journals tout their IF, like this example) is inherently problematic. Hypothetically, a journal which published many highly criticized articles could achieve a significant IF through the negative discussion of its works!

Secondly, there are some questions about data integrity. When a citation to a work is indexed, no human checks are being done to ensure the validity of that citation, for example, that the paper exists and there no errors.

Thirdly, many critics have claimed that WOS/IF exhibit a bias towards English-language and perhaps primarily Western journals—which might be true, at least insomuch as some
research disciplines see this same bias, and WOS/IF reflects the scholarship being most heavily discussed in those disciplines.

Fourthly, IF is subject to manipulation. Any evaluation done through citations alone is subject to a certain amount of “gaming the system” through author self-citations and reciprocal citation agreements between authors. But the other issue of manipulation that arises with IF is that of coercive editorial practices. This may take the form of a journal editor encouraging, pressuring, or obligating a submitting author to cite recent articles from that same journal. It may also take the form of an editor seeking to condense some submissions by removing abstracts, reducing references, or similar to convert them into items they think will not be deemed “citable,” attempting to improve the journal’s overall ratio at the expense of the scholarly work and its value to serious readers. Some critics have also asserted that the priority on IF compels some authors to butcher their own work to suit the format and style of high-impact, broad-audience journals (versus lower-impact journals targeted at specialists), even if this leads to gross oversimplification of findings and conclusions (Werner 2015).
Earlier I mentioned that IF is also provided in a 5-year ratio. This sometimes offers a useful alternative to the traditional 2-year IF, because it may better address the natural and expected variation in IF’s over time—this variation can be as high as ±15% per year for large journals, and as much as ±40% per year for smaller journals (Engle & Walker). The 5-year IF can be found in *Journal Citation Reports* (JCR) along with the 2-year IF—JCR is available through the SHSU Library’s website.

The Eigenfactor is built on the same proprietary data from Clarivate Analytics that is used in calculating IF. Where it differs is that Eigenfactor gives different weights to individual citations, asserting that not all citations are created equal. Citations from higher-prestige journals are more highly valued than citations from lower-prestige journals. In situations where journal-level metrics are appropriate, including an Eigenfactor alongside 2-year and 5-year IF’s may provide a more three-dimensional picture, not only of how many citations the journals receive, but where they come from. Eigenfactor is now included in JCR but can also be freely obtained at eigenfactor.org

Scopus Journal Rankings is a suite of metrics from SCImago. These metrics are
based on data from Elsevier’s Scopus product, which is a competitor to Clarivate’s Web of Science, and thus the index of journals on which calculations are based is different. Additionally, SJR creates rankings using an algorithm, based on Google’s PageRank, which calculates “article influence” (Roemer and Borchardt 2012). Because it is calculated from a different database of journals, SJR may rank certain journals that do not have Impact Factors, or may simply provide an alternative perspective of a journal alongside its IF. SJR may be freely obtained at SCImagoJR.com.

A metric closely related to SJR, and also based on the Scopus database, is SNIP, the Source-Normalized Impact per Paper. SNIP “corrects for differences in citation practices between scientific fields, thereby allowing for more accurate between-field comparisons of citation impact” (journalindicators.com). SNIP and related indicators are freely available from CWTS at journalindicators.com – CWTS even calculates a stability interval to indicate the reliability of a journal’s SNIP.

Cabell’s Scholarly Analytics is a resource provided by the SHSU Library that indexes and reviews journals in a variety of disciplines (though its coverage is not comprehensive). They display IF for journals when available, but they also calculate their own “Cabell’s Classification Index” in an attempt to quantify “how much influence a journal has in your field of study” (Cabells’). The screenshot illustrates a journal’s CCI score for its discipline and the specific topics that it covers.
In addition to these other citation-based metrics, other considerations for evaluating journals could include acceptance rate (how hard is it to get published with them?) and subscription and circulation rates (which provides at least some rough gauge of readership). Acceptance rates can often be found through tools such as Cabell’s Scholarly Analytics or the MLA Directory of Periodicals, and circulation rates are also sometimes available through MLA Directory of Periodicals. If these numbers are not readily available for a specific journal, try calling the subscriptions or public relations office of a journal to request this data.

The first screenshot comes from Cabell’s and shows some of a journal’s details, including acceptance rate.

The second screenshot comes from the MLA Directory of Periodicals and shows a journal’s circulation, as well as the numbers of articles submitted and published each year (some quick math indicates a 16.7% acceptance rate).
We also have the option to look more granularly at a specific article, rather than the journal where it appeared.

The most obvious of these is the actual citations to a work. But more useful than a simple count is citation context analysis: providing insight into who has cited the article, in what publication(s), to what extent, and for what purpose(s). For instance, a passing reference to an article in a literature review, where it shares a footnote with four other studies, is less noteworthy than an article receiving credit for establishing a new methodology that is being reused by other researchers. Data on citing works can be gathered via library research databases like Web of Science, free tools like Google Scholar or Harzing’s Publish or Perish software, researcher networks such as Mendeley and ResearchGate, and other types of strategic searches.

Be sure to have a sense going in of whether you want to focus only on citations in scholarly works or whether you also want to include citations in gray literature such as government, industry, and non-profit reports, patents, clinical trials, white papers, conference presentations, and so forth, as these other types of citations may have greater or lesser value depending on the researcher, the discipline, and
the goal of the published works. Researchers particularly interested in identifying these other types of citations are also encouraged to leverage search engines such as GreyNet, PapersFIRST, ProceedingsFIRST, Google Patent Search, USA.gov, THOMAS, the National Technical Information Service (NTIS), the Educational Research Information Center (ERIC), the National Criminal Justice Reference Service (NCJRS) database, the World Health Organization (WHO), or others.

Discussion of a work in traditional media and press outlets may also be valuable for some researchers. You may be aware of these if a journalist contacted you, or they will often be easily turned up in a standard Google.com search or a strategic search in the SHSU Library’s newspaper databases.
Informal and organic discussion, shares, or “likes” of scholarship on social media channels may provide evidence of a research work’s impact on the field. However, these metrics can also be more challenging and time-consuming to collect and to interpret—for instance, there is little agreement industry-wide about what Twitter buzz really “means,” so you need to be prepared to explain these numbers in the context of your field.

Some PlumX altmetrics are available by searching for an article in the SHSU Library’s Engine Orange search tool. A free profile with ImpactStory will compile some of this data. The commercial tools Altmetrics and Plum Analytics also provide free tools to obtain a snapshot of social media activity around a specific work, as long as it has a Digital Object Identifier (DOI). All of these tools will be linked in the presentation slides for your future reference.

Many researchers already participate in broad academic networks, such as ResearchGate, Academia.edu, or Mendeley; more discipline-focused networks such as Social Science Research Network (SSRN); or even niche networks such as BioMedExperts. These networks provide yet another avenue for documenting engagement of and influence on other researchers in the field. Many such sites will document who views or bookmarks a researcher’s profile and citations, and they
may additionally track views and downloads of individual articles, if a researcher shares an article on the network as permitted in their publication agreement. Because Mendeley at its core is a reference management tool, it can also provide unique metrics about how many other users have added your work to their library—this indicator often correlates positively to future scholarly citations.

Additional data on document views, downloads, and web traffic may be available from the journal publisher, as illustrated in this author dashboard from Taylor & Francis.
Even more metrics and quantitative approaches exist to quantify researcher productivity and impact, gauge institution-wide research output, and measure the impact of books, interdisciplinary works, and more. I have tried to highlight key alternatives to IF for assessing the impact of a journal or a specific article, but I have barely scratched the surface of bibliometrics and scientometrics. If you want to learn more about additional metrics, I hope you will refer to the selected references I have included at the end of this presentation.
But before we come to those references, I do want to share a few thoughts about the limitations of metrics in general, because we should understand both the strengths and the shortcomings when we engage with these assessment tools.

First of all, it pays to maintain perspective on any metrics based in citations, because it isn’t always clear why authors cite and what the inclusion—or exclusion—of a citation is really meant to communicate.

“Enthusiasm for citation indexing...is based on the assumption that citations give a fair picture of the intellectual links between publications. It would be more accurate to say that they give the picture that authors record. The deviation results from memory failures, lack of self-awareness, carelessness, plagiarism of other people’s citations without having actually used them, the widespread custom of not citing ‘obvious’ sources, and many other causes—all consequences of the simple fact that the author selects citations to serve his scientific, political, and personal goals and not to describe his intellectual ancestry.” – Kenneth May

But we should also acknowledge the limitations of research metrics more broadly.

“All research assessments (whether using metrics or peer review) are essentially
backward-looking and based on past performance. We assume that because someone or something has performed in the past, it will do so in the future. But that is an assumption.” (Gadd 2018)

I already identified issues of potential manipulation of IF, and in truth, “gaming the system” usually just requires motivation. Altmetrics such as views or downloads could be manipulated just as easily, if not more easily, than citation-based metrics. Any number can be used out of context, misconstrued, or misinterpreted. And sometimes there are gaps in numbers, but the absence of readily available metrics does not automatically equate to the absence of impact, a fact of which I frequently try to reassure individual researchers.

When faculty feel pressured to publish only in high-impact journals, which tend to be very expensive, who is being excluded? Researchers at small institutions with smaller library budgets, researchers in developing countries, and others will not have access to the work in these scholarly journals, so they are excluded from the conversation we create.

Fundamentally, there is also an argument that the over-emphasis on arbitrary measurements to justify research as “worthwhile” undermines the philosophy that the pursuit of knowledge has intrinsic value. If a work is not cited within 2 years – or within 20 years – does that indicate that the work has no value? That it does not have the capacity for future value? That the researcher is not still worthy of recognition for their pursuits and accomplishments?

The 2012 San Francisco Declaration on Research Assessment (DORA, https://sfdora.org/) actually recommends the abandonment of the use of journal-level metrics in hiring, promotion, and funding. Nearly 600 academic and scholarly organizations have signed on in support so far, along with nearly 13,000 individual scholars. This document represents a dream of a scholarly world based entirely in formative, qualitative assessment rather than purely quantitative benchmarking.
BUT – Like it or not – For the moment, research metrics are part of our reality. Sometimes scoring and ranking are necessary because there can be only one selected candidate for a job, only one grant recipient. So my recommendation is to approach metrics with caution but also with practicality. Gather and present research metrics as you feel is warranted in grant applications, tenure and promotion packets, CVs, and so forth, but seek to follow some best practices along the way.

• Provide multiple, varied metrics whenever possible. Do not rely on any single number in isolation.
• Select the most relevant and appropriate metrics for your discipline, project, and purpose. Your array may vary from another researcher’s.
• And seek to place metrics in context, presenting a cohesive narrative of your work and its place in your discipline and the larger world.
The other available approach is to work towards changing the conversation around research metrics. Open a dialogue within your department, your professional association, or your field more broadly to build a consensus about

- Which metrics are most appropriate for your field and for the medium in which you disseminate your research;
- How it is or is not appropriate to apply those metrics in evaluation, tenure, and promotion decisions;
- What role formative evaluation (feedback, peer review, qualitative assessment) should play vs summative evaluation (metrics, quantitative assessment)?

Champion the development of a select journal list in your department—rather than relying on external quantitative metrics, leverage the subject expertise in your department to identify preferred high-quality journals in which faculty should focus on publishing. I’m aware of some departments in the College of Education who have done this, and I think it can be more effective than trying to set targets for metrics like IF. But do recognize that such a select list should be revisited to a regular schedule and revised as necessary to accommodate new or emerging journals, a journal of declining quality, etc.

Finally, work to foster a culture of recognition rather than a culture of evaluation, which is comparative, competitive, and often correlated with mental health issues in academia. Set
minimum expected standards that everyone should meet, yes, but beyond that, practice recognition for the different strengths and accomplishments that diverse faculty bring to the table.
SELECT REFERENCES / FOR FURTHER READING

- Blaise Cronin & Cassidy R. Sugimoto (eds.), Scholarly Metrics Under the Microscope: From Citation Analysis to Academic Auditing (Information Today, 2015) - An incredible read that draws together 60 years worth of both classic and recent papers on metrics
- San Francisco Declaration on Research Assessment (2012. https://sfdora.org/)
LINKS TO SELECT RESOURCES / TOOLS

- Through the SHSU Library:
  - Journal Citation Reports
  - Web of Science
  - Cabell’s Analytics
  - MLA Directory of Periodicals
  - PapersFIRST
  - ProceedingsFIRST
  - Engine Orange (find PlumX altmetrics)

- Online:
  - Eigenfactor.org
  - SCImago SJR
  - Journal Indicators (SNIP, etc.)
  - Google Scholar
  - GreyNet
  - Harring’s Publish or Perish free software
  - Mendeley

- ImpactStory
- Altmetric Bookmarklet
  (article altmetrics)
- PlumX DOI lookup
  (article altmetrics; add your DOI):
  https://plu.mx/altmetrics/
I developed and maintain the library’s Guide to Scholarly Writing, Publishing, and Research Impact. There you can find additional details on the metrics I discussed today and resources on many other topics in the realm of creating and disseminating scholarly work (scholarly communications). I will be posting the slides from today’s presentation, complete with my presenter notes, if you wish to refer back to anything.

Reach out to me to request one-on-one consultations or training in finding and gathering metrics or understanding many other topics and tools in scholarly communications.

Schedule a presentation for any group on any topics of interest in the realm of scholarly research, writing, publishing, impact, open access, or any other aspect of the scholarly communications lifecycle.
Promote data management session in Spring 2019
Comments? Questions?

Erin Owens
Associate Professor, Newton Gresham Library
Access Services Coordinator & Scholarly Communications Librarian
936-294-4567
eowens@shsu.edu

My ORCID Researcher ID: 0000-0001-9520-9314