

Journal Impact Factors

August 30, 2013



Every summer my e-mail is enlivened by people and organizations writing about the latest journal impact factors (IF). Because I chair the APS Publications Committee, I have always done my best to feign deep interest about IFs. I know many people take them very seriously, but the truth is, I have never cared about them too much, although I do look at them.

I certainly take citations seriously; they indicate at least to some degree the impact and worth of a paper. But journal impact factors seem to have been created by people who were a bit shaky on both arithmetic and descriptive statistics (the kind you learn in the first two weeks of baby stats classes in psychology).

First, the arithmetic problem. IFs always go to three decimal places, to thousandths of a citation, giving a spurious impression of precision. The most recent IF for *Psychological Science* (*PS*) is 4.543. Consider, however, its derivation: The number of citations in 2012 to articles published in the previous two years (a whole number, of course — it was 2344 for *PS*), divided by the number of articles published in *PS* in those two prior years (another whole number — 516). So dividing 2344 by 516 gives the IF. But by the rules of rounding we all learned somewhere around fourth grade — round to one decimal place more than the raw data — the proper IF is 4.5. Rounding to two places is sometimes done — 4.54 in this case — but why would anyone ever round to three decimal places — to thousandths — after dividing one whole number by another one? I have no answer. Why not, by this algorithm, use 4.5426356 as IF? After all, my calculator gives me this many digits to the right of the decimal point when I divide 2344 by 516.

In many fields, journal IFs do not differ much among journals or from year to year, so using three decimal places also gives an impression of change where little or none exists. For example, the IF for *Psychological Science* “skyrocketed” from 4.431 in 2011 to 4.543 in 2012 (or from 4.4 to 4.5 when rounded appropriately). Rounded to one decimal place, the humdrum constancy of the IF for most

journals I have examined produces ennui. They just don't change much. Yet when they do, even by a tiny bit, people get excited.



At a Glance

$$\text{Impact factor} = \frac{\text{Citations in 2012 to articles published in 2010 and 2011}}{\text{Articles published in 2010 and 2011}}$$

The first e-mail report I received about a journal IF this year was for the *History of Psychology*. The correspondent reported that she had received “fantastic news” because “the 2012 impact factors have been released, and ... *History of Psychology* has received a rating of .750. This is an increase from the 2011 impact factor of .265.” So, the IF is still well under 1.0 — tiny — but nonetheless there is reason to celebrate! Actually, I think it is unfair to judge journals like *History of Psychology* by IF; other criteria, albeit harder to measure, are more relevant for this and many other journals — use of articles in textbooks and courses, for example. It's also more accurate to refer to the IF as a mean, not a rating.

Carrying to three decimal places is the least of the problems with the IF. More critical is the fact that its creators apparently missed the first two weeks of Basic Statistics: If a distribution is strongly skewed, the mean of the distribution provides a measure of central tendency that can be greatly affected by some dramatic outliers. Of course, the IF represents the mean impact. I have been examining (and sometimes collecting) citations for one reason or another for 35 years, and their distributions are always staggeringly skewed in a positive direction. Most scientific papers are not cited much, but a few are cited a huge number of times. Plot the distribution and it will always be strongly skewed to the right (i.e., the bulk of the numbers fall on the left, often between 0 and 3 on a yearly basis). The tail of the distribution accounts for most of the variability in IF among journals. Thus, when one considers a change in the IF across years, this may be due to the presence (or absence) of just a few highly cited papers in the last two years, not an overall shift in the distribution. For *Psychological Science* in 2012, the median impact score is 3.0 and the mode is 0.0. (Yes, the most frequent number of citations for the articles published in 2010 and 2011 was zero — for about 15 percent of the papers.) The highest cited paper was the “false positive” paper by Simmons, Nelson, and Simonsohn published in 2011 (69 citations). Perhaps it caused the uptick in IF for 2012; the next highest paper was from 2010 and was cited 27 times.

As an aside, not only does this point hold true when examining journal articles, but the same tenet holds for citations of individual researchers: Some researchers are cited a lot, others not much at all. Part of the issue is that different fields have very different citation patterns. The same point even holds true for researchers who are very highly cited: Most papers, even of highly cited researchers, are not cited very often. The researchers become “highly cited” because a small subset of their papers are cited at very high rates.

A case could probably be made for the mean or the median for impact factors, at least for journals that publish a large number of papers. Another measure might be the number of papers with greater than X citations (where reasonable individuals can disagree on what X should be). Eugene Garfield, creator of the original Institute for Scientific Information Citation Index, deemed that a paper cited 100 times should be considered a “Citation Classic.” However, citations vary widely by fields and even within

subfields in psychology, so there is no one-size-fits-all answer to what a “classic” should be. One hundred is probably as good a number as any other for psychology, although of course the number of years since publication should be taken into account (a paper that reaches 100 in three years is different from one that takes 12 years). Of course, scholars in the field of scientometrics debate about various ways to measure impact of individuals, with the h-index the most commonly used (see my column “[The h-index in Science: A New Measure of Scholarly Contribution](#)” in the April 2006 issue of the *Observer*, although it is a bit dated now). In addition, researchers have devised other ways of measuring journal impact, the most important being [Eigenfactor scores](#), perhaps the topic for a different column.

Unpacking Impact

Because of the way IFs are calculated, total impact of journals is completely missed. Let’s consider the top two general psychology journals in terms of impact factor for 2012. These are almost always the same two or three titles, which are review/theory (versus empirical) journals. For 2012 they are *Psychological Bulletin* (15.575) and *Annual Review of Psychology* (15.265), so by this measure the impact of these two is about the same. They finished in a dead heat (many years *Annual Review* is number one). But really? Let’s look at total citations in 2012 instead. The total citations for *Psychological Bulletin* were 30,814; those in *Annual Review* were cited “only” 10,635 times (still quite good). Which outlet actually has the greater overall impact on the field? Why not look at the total citations of a journal? *The Annual Review of Psychology* is a great publication, but it is, in some ways, an odd one to include as a “journal.” It is published once a year, the editorial team handpicks the review topics and the authors, and they often try to commission reviews of hot, breaking-news topics. I served on the editorial board for six years and saw the process in action. It is more like an edited book than a journal, where for the latter most anyone can submit a paper. Because the chapters are handpicked and written by leaders in the field, it would be a wonder if *Annual Review* chapters were not highly cited.

APS Journals Are in the Top 10 Percent of IFs

In the great ocean of psychology publishing, APS journals have always held their own. Citation figures exist for three of our journals, and Thomson Reuters puts them in the same category, one called “Multidisciplinary Psychology,” which contains 126 journals. The grouping of journals seems another example of poor judgment by someone, because our rated journals — *Perspectives on Psychological Science* (ranked 5th at 6.594), *Psychological Science* (8th, 4.543), and *Current Directions in Psychological Science* (10th, 4.222) — are lumped in with a wide variety of journals that do not seem much like the first 10 journals in the list (which mostly cover all of psychology). Others in the total list include *Body Image* (37th), the *Journal of Gambling Studies* (41st), *Dreaming* (76th), and *Humor* (101st), among many others. These might all be fine journals, but they seem rather specialized to belong in the same category with *Psychological Bulletin*, *Psychological Science*, and other journals that are truly general. Nonetheless, whether the category of “Multidisciplinary Psychology Journals” makes sense or not, the three ranked APS journals are all in the top 10 percent in terms of IF. In terms of total impact in 2012, *Psychological Science* led the way with 17,503 total citations, with *Current Directions* having had 5,016 and *Perspectives*, 2,227. Of course, that ordering also is the same as the number of articles published per year in the journals. -H. L. R., III

A related problem with IF is the vast differences in the denominators across journals. Even for empirical journals, some contain many short articles each year, whereas others contain fewer but longer articles. Because of the highly skewed distribution of citations noted above, a single high-citation paper will have more of an influence on the IF when the denominator is small. It is difficult to know how much this matters, because of course the

“many article” journals have more opportunities for a big hit. Still, it would be interesting to calculate IF on the basis of (say) number of citations per 10,000 words of print in the journal. This step would also take into account the difference in total pages printed across journals.

Another issue in considering IF is the size of the readership. For example, *Psychological Science* is sent out to all members of the Association for Psychological Science, so the IF may naturally be skewed by the size of the readership, as opposed to the quality of the papers. Still, the fact that all our journals go to all our members is a strong reason to want to publish in APS journals even if they had low impact factors (and they don't). The APS journals currently go to all 25,000 members (electronically or on paper), providing a huge potential readership relative to virtually all other psychology journals.

I have heard people say they want to submit to high-impact journals, as indexed by IF. I suppose that makes some sense, but I wonder if there is any evidence to support the idea that an article will be more highly cited if it is in one place or another? Of course, the only real way to test this would be to publish the same article in both venues, so by definition, it is impossible to ever gain evidence on this point.

Impact Factors of Psychology Journals

Psychology journals have low impact factors relative to journals in many other scientific fields or relative to more general science journals. Consider *Nature*, *Science*, and *Proceedings of the National Academy of Sciences*, with 2012 IFs of 38.6, 31.0, and 9.7, respectively. A quick look at these IFs makes it clear that if a journal's IF is 4.431 one year and 4.543 the next, it's really pretty stable and there is no cause for celebration in the increase; likewise, a small drop is no cause for concern.

A variety of reasons exist for psychology's relatively low IF compared with other fields (in neuroscience, for example). Our publication processes are slow relative to other fields; often editors give reviewers two months to review a paper, multiple rounds of revision are required, and then even when the paper is accepted the lag between acceptance and publication is much longer than in other fields.

So if a paper in my field comes out this week and I cite it in an article I am submitting next week, given all the factors above, the paper might not even appear in the journal in two years (assuming it is accepted). And the pace of research is just slower in some areas than others due to the types of research conducted. -H. L. R., III

Given all of the above

concerns (and others), I am always surprised to hear people say that IFs are (or should be) used in promotion and tenure judgments — is the candidate publishing in high-impact journals? I have also heard that journal IFs may be used in making funding decisions. To me, these are sad misuses of the IF. The critical question is “What is the quality and the impact of the candidate's work?” for a tenure or promotion decision. That might be correlated (albeit imperfectly) with the IFs of the journals in which the candidate publishes, but the focus should be squarely on the candidate and her/his work, not on correlated factors like journals where the candidate publishes. The question is “Does the candidate publish in appropriate journals for his/her field?” Most scientists work in specialized fields and hence publish in specialty journals.

The *Journal of Experimental Psychology: Learning, Memory and Cognition (JEP:LMC)* is one in my own field. Its impact factor is 2.918 (18th in the list of experimental journals in psychology, according to Thomson Reuters). That's not up there with *Psychological Bulletin*, but of course it is an empirical (specialty) journal and their impact factors are never as high as those for review journals. Still, if I have what I consider a good empirical paper, I would not hesitate to submit it there. A paper I published in *JEP:LMC* with Kathleen McDermott in 1995 has been cited 2,134 times (according to Google Scholar),

or 118.6 times a year. On the other hand, I have another first-authored paper in *JEP:LMC* from 1983 that has been cited 1.3 times a year. The latter, sadly, is more typical. The point is that it is the paper, not the journal, which counts for evaluating individual papers. The same goes for a person's body of work. Some fields simply do not lend themselves to high citation counts (psychophysics, animal learning, and history of psychology come to mind). Measures of impact should be relative to those in the field in which the researcher works.

The Devil You Know

I have been hard on IF, but let me play devil's advocate and admit that IF, even with its flaws, captures something real about journal quality. Look at any list created by Thomson Reuters in your field and you will doubtless conclude as I have that the top 15 percent of the journals are better than the bottom 15 percent. If I were on a university promotion and tenure committee and a candidate in, say, political science, was presented by her chair as publishing in the "best journals in the field," I could check the IF values of political science journals and see for myself. As noted above, if the journals in which the candidate is publishing are not among the top as measured by IF, there may be a perfectly valid explanation; the candidate may be publishing in the top journals of her subfield. Even if IF values have some uses, when judging job candidates, tenure candidates, and grant proposals, the focus should be on the individual (or proposal), not just on IF of journals where the candidate published.

What IFs

There is much more to be said about IFs, and I have discovered that many people besides me are saying it. Most are taking a dim view. Just Google around; also read the article in Wikipedia and follow its links. APS has signed the *San Francisco Declaration on Research Assessment* (subtitled *Putting science into the assessment of research*). This statement grew out of a meeting of the American Society for Cell Biology in San Francisco in December 2012. It notes that journal impact factors were originated primarily to help librarians decide which journals to purchase, but that now the IF is being used for all sorts of other purposes (tenure decisions, funding decisions). The *San Francisco Declaration's* first General Recommendation is "Do not use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist's contributions, or in hiring, promotion, or funding decisions." It then goes on to make more specific recommendations about how to appropriately assess the individual researcher's contribution, focusing on her or his body of work in relation to others in her or his specific field. The entire statement can be found on [DORA's website](#).

APS is one of the 82 original signers, with the total number now at 329 organizations as of this writing. Assuming this trend catches on, my e-mail in late June and early July every year about the new IF data may lessen.