

Journal-level metrics are poor surrogates for research evaluation The variety of variants to the Journal Impact Factor might at first seem to be an indication that there is a lot of competition to win the favor of librarians. Because, of course, journal-level impact metrics like the Journal Impact Factor were born of the desire to rank the importance of journals so that librarians will know which ones to have in their libraries.

And while modern journal-impact bibliometrics may be only distantly related to the original work by Gross and Gross (1927), they have the same focus: ranking the journals themselves. The leap seems to be huge, from journal rankings to judgements of the quality of a individual articles in the journals, but in practice, it's often unnoticed. For example, in the introduction to an article describing yet another variant on the Journal Impact Factor, the authors write:

“The citedness of a scientific agent has for decades been regarded as an indicator of its scientific impact, and used to position it relative to other agents in the web of scholarly communications. In particular, various metrics based on citation counts have been developed to evaluate the impact of scholarly journals....” (González-Pereira *et al.* 2010)

In other words, since the number of citations received by a researcher, or research group, can be a useful indicator of the quality of their research, it is useful to count the number of citations received by the ~~researchers~~ journals in which the researchers publish. The logic of which seems to be reversed: instead of extrapolating that good articles make the journals they are in better, we extrapolate that good journals somehow make the articles that are in them better.

As Adler *et al.* put it:

“...Instead of relying on the actual count of citations to compare individual papers, people frequently substitute the impact factor of the journals in which the papers appear. They believe that higher impact factors must mean higher citation counts. But this is often *not* the case! This is a pervasive misuse of statistics that needs to be challenged whenever and wherever it occurs.” (Adler *et al.* 2009)

Yet, it's hard to explain the variety of bibliometrics available to assess the impact of journals, and the nature of some of the refinements that differentiate them [[EXAMPLES?]], without assuming that the metrics are being used for something more generally desirable than the ranking of the journals themselves.

At the same time, however, it is this “pervasive misuse” of journal-level metrics that has helped lead to the development of article-level metrics. (See below.)

References

- Adler, Robert, John Ewing, and Peter Taylor. 2009. “Citation Statistics.” *Statistical Science* 24 (1): 1–14. [doi:10.1214/09-STS285](https://doi.org/10.1214/09-STS285).
- Gross, P. L. K., and E. M. Gross. 1927. “College Libraries and Chemical Education.” *Science* 66 (1713): 385–89.
- González-Pereira, Borja, Vicente P. Guerrero-Bote, and Félix Moya-Anegón. 2010. “A New Approach to the Metric of Journals’ Scientific Prestige: The SJR Indicator.” *Journal of Informetrics* 4 (3): 379–91. [doi:10.1016/j.joi.2010.03.002](https://doi.org/10.1016/j.joi.2010.03.002).