https://doi.org/10.51408/issi2025_152

Trueblood et al.'s Ideas on Research Evaluation and Implications for Reforming Research Assessment

Ronald Rousseau

ronald.rousseau@uantwerpen.be, ronald.rousseau@kuleuven.be
University of Antwerp, Faculty of Social Sciences, 2020 Antwerp (Belgium)
KU Leuven, MSI, Facultair Onderzoekscentrum ECOOM, 3000 Leuven (Belgium)

Abstract

This contribution to the FRAME track at ISSI 2025 offers a brief overview of Trueblood et al. (2025), highlighting its relevance for research evaluation. In their article, Trueblood and 14 co-authors examine the current publication landscape and explore both how it can be transformed and how such changes logically necessitate a shift in research assessment practices.

Introduction

Early this year, Trueblood and 14 co-authors published "The misalignment of incentives in academic publishing and implications for journal reform" in the Proceedings of the National Academy of Sciences of the USA (PNAS). The authors studied the publishing landscape, decry how commercial publishing companies generate huge profits, and propose ways to let academic institutions regain control of scientific publishing. As scientific publications play an important role in research evaluations, the authors spend several pages on reforms in academic evaluation. In this contribution to the FRAME track at ISSI 2025, a short overview of Trueblood et al. (2025) is presented with an emphasis on its implications for research evaluation. In this document the expression "the authors" is used for Trueblood et al. (2025).
The authors noted that the two main goals of publishing, namely the documentation of new knowledge and establishing scientific credentials are often in tension. It is, indeed, well-known that even in the best of circumstances, maximizing metrics may lead scientists to prioritize novelty and even sensationalize findings to publish in so-called prestigious journals. In this way, important details and partial null results may be hidden from view.

Academic publishing: now and in the past

The development of academic publishing is closely connected to the growth of universities, the formation of scientific societies, and the professionalization of academia. Trueblood et al. (2025) write that in 1950 there were about 10,000 journals worldwide. This number increased to 62,000 in 1980 and according to (Suiter and Sarli, 2019) to 80,000 in 2019. Note that in 2015 an estimate for the number of Chinese journals has been published, reaching a total of at least 8,000 journals (Rousseau, 2015). Commercial publishing companies such as Elsevier, Springer Nature, Wiley-Blackwell, and Taylor & Francis, dominate a large part of modern scientific publishing. It is sometimes argued (Fyfe et al., 2017) that these firms

exploit reviewers and editorial boards by requiring free services, making it costly to distribute scientific work, and levying high fees for open access.
Trueblood et al. (2025) further discuss the role of modern journals. They offer details from three perspectives: 1) journals as revenue streams, including the bad sides of it such as predatory journals and paper mills; 2) journals as curators of research and the role of peer review; 3) journals as the cornerstone of the academic prestige economy, leading to problems such as "publish or perish" and the pressure to continuously produce and publish scholarly work, preferably in high-prestige journals.

Alternative publishing models

The authors argue that academic institutions and learned societies should take over the journal publishing industry, turning it into a nonprofit sector, where science controls science. New publishing models, including preprint platforms, must be established. They propose some ideas and offer examples of existing initiatives.
- Academia retaking control
As commercial publishers are not likely to stop their business or hand it over to academia, the following steps are proposed. First, academic institutions and associations create new journals controlled by themselves and ask editorial teams now working for for-profit journals to switch to the new journals. Second, scientific academies and societies should support this switch and ask their members to stop working for commercial companies, either as editors, reviewers or as authors. Third, as academics cannot take care of the purely technical work of journal publishing, competitive calls must be made so that experienced companies can do this work at lower costs. Note that this comes close to what ISSI has done when switching from Elsevier to MIT Press. In conclusion: publishing must be by scientists for scientists.
- Preprint servers
In certain fields preprints and society proceedings are considered already as more reputable than journal publications. This has happened because highly cited articles may reside solely on preprint servers. Perelman's Fields Medal winning mathematical work has never been formally published (he even refused the Medal). In our field we have (Larivière et al., 2016), cited 220 according to Google Scholar, but never formally published in a journal. Generally, and in all fields, preprints are becoming increasingly valuable. Besides citation counts, also download counts are calculated for preprint papers.
- Journal reviewed preprints
This section answers the problem that preprint servers contain papers that are not peer-reviewed. Here Trueblood et al. (2025) refer to the new policy adopted by eLife. Nowadays this journal only reviews articles made available on a dedicated preprint server. The outcome of these reviews is no longer used as a basis for an accept/reject decision, and the number of articles hosted on eLife is not limited. The decision of whether to host a submission, before review, is made by scientifically active editors, and reviews are presented as commentaries alongside the article. Although the authors consider this approach a significant change, they do not call it "disruptive", as editors still determine which papers will be hosted on the preprint server. The idea

of journal reviewed preprints can focus the publishing process on improving reporting and facilitating knowledge diffusion.
Yet, because the method differs so much from the traditional way, Clarivate has decided that eLife will not receive an impact factor in 2025.
- Community reviewed preprints
PCI (Per Community In) also undertakes to review preprints. PCI is a community-sourced service that provides free, journal-independent reviews of preprints. Preprint authors and reviewers collaborate to improve the preprint, ultimately leading to a recommendation where a recommender (serving a similar role as a journal editor) endorses the article for publication. The process may conclude when a PCI-recommended preprint is published on the corresponding thematic PCI websites with a DOI, allowing it to be cited, or it may be published directly in a PCI-friendly journal. PCI and similar initiatives align perfectly with an open access framework. The concept of community-reviewed preprints bears similarities to the idea proposed by Perakakis et al. (2010) under the name of Natural Selection of Academic Papers. Instead of a service like PCI, they envisioned a Global Open Archive containing the original preprint (possibly via an institutional preprint server), open and signed reviews (including those initiated by the authors), updates, and citations. In this open environment, also reviewers could be rated.
- Society endorsed preprints
The role of PCI could be assumed by societies, enlarging the prestige of this approach. The authors suggest that also ArXiv, and similar preprint servers, could play this role. They further propose that federal granting agencies and private foundations could supply the resources needed to support these changes. However, they note that this approach still carries the notion of prestige, much like traditional journals. As a result, they conclude the section by suggesting that perhaps scientists should move away from using publications as a measure of prestige.
- Modular publishing platforms and micropublications
Modular publishing breaks up a paper into small sections called modules. According to the authors F1000, eLife and PLOS Biology already publish micropublications, small articles without a broader context. Some platforms such as Octopus allow the threading of modules into a coherent narrative.

Barriers to change

The more new publishing models differ from the classical way publications are handled, the more difficult it is to become broadly accepted. There is a monetary cost and the cost of extra learning and effort, deterring potential adopters.
Pay-to-publish is an obvious choice for these new models, but requires that scientists have funds at their disposal and micropublications do not seem to correspond with the way some fields, such as the humanities see scholarly work.
Trueblood et al (2025) end this part with reflections on barriers to change by stating the following three big challenges to journal reform:

a) The lack of independence of most scientific journals from commercial for-profit publishing companies

b) The financial impacts on societies that nowadays generate substantial revenue from their journals

c) Resistance to adoption because of concerns regarding academic prestige

This leads to the topic of this conference section.

Reforms in academic evaluation

For most scientists today, publishing needs to translate into career value, namely, recognition by hiring committees and funding agencies. Therefore, the reputation of a publication venue is crucial. This highlights the need to reform academia's incentive structure, but the authors caution that such changes could bring unintended consequences. They point out that the current system has already contributed to the rise of paper mills and so-called 'predatory journals'. To better understand these potential pitfalls, they suggest that applying game theory could offer valuable theoretical insights.

Next, the authors consider five ways in which to reform academic evaluation.

Abandoning problematic metrics

They recall that citation counts and journal impact factors are highly problematic. The number of received citations depends on many factors, many of which are independent of research quality (including pure luck). Since Seglen (1989, 1997) and its replications (Zhang et al., 2017), researchers know that impact factors should not be used to judge papers. This insight has led to many reactions from the scientific and publishing community such as, e.g., DORA (2012). Moreover, altmetrics are even more easily gamed than citations. As examples of positive evolutions, the authors mention the introduction of narratives and evidence-based curriculum vitae. Of course, an academic career path does not only include published articles, but also books, teaching, and outreach activities.

Adopting responsible metrics

Here the authors mention the Leiden Manifesto for Research Metrics (Hicks et al., 2015), emphasizing best practices and allowing researchers to hold their evaluators to account. They further recall that using evidence-based CVs is an example of a best practice, as it leads to transparency. Yet, one must recognize that there are at the moment systemic issues in the evaluation process.

Quantitative metrics: measuring researcher impact

In this section, Trueblood et al. (2025) consider the evaluation of researchers and discuss different ways of weighing authors' contributions. They do not come to a concrete proposal for this very tricky problem. In some counting systems, for an overview of counting systems we refer to (Gauffriau,2021) adding authors decreases the score of the others, which could result in scientists with disadvantaged backgrounds being relegated to the acknowledgment section.

Quantitative metrics: counting replications

The push for novelty makes researchers reluctant to try to replicate others' work. Yet, it is well-known that many published results cannot be replicated, the so-called "replication crisis". It is suggested that the number of replications (and I add also direct extensions) could be a measure of interest created in the field (or even outside). If successful, replications, and replications of replications can be a measure of the robustness of the original research. Note that the authors even include unsuccessful replications, of course not in terms of robustness but in terms of research interest. They, correctly, warn that emphasis on replications and reproducibility should not divert scarce resources.

Rewarding societal impact

Since much of the research is funded by public institutions, it's reasonable for the public to expect some return on that investment. In this context, scientists are expected to engage with the broader community and, ideally, to make a visible impact. While measuring the outcomes of such engagement can be challenging, a useful starting point might be tracking the input, namely, how often scientists interact with the public. Talking about the societal impact the authors refer to Overton.io for the quantification of the policy influence of publications and of grey literature such as technical reports.

Incentivizing quality over quantity

The authors point out that there are some easy and relatively minor changes possible in academic evaluation to alter too narrow incentive structures. Focusing on top x publications in the latest y years is such a simple measure. This alters the focus from quantity to quality. Of course, it is supposed that evaluators actually read these papers, otherwise, the focus would shift again to "high-level journals", or worse to impact factors. The authors provide examples of funding organizations that take this approach such as ERC (Europe) and the NSF in the USA. Papers must be separated from the journals in which they are published. This is a must when papers are not published in the traditional sense of the word, cf. the earlier section on publishing.

Discussion and Conclusions

The influence of commercial publishers and the academic prestige economy have both a detrimental influence on scientific quality and the idea of science for the benefit of humankind. Biased incentives have even led to academic fraud such as using paper mills to increase the number of publications and to other fraudulent behavior.

The authors propose that publishing goals should be aligned with the broader aims of knowledge creation and dissemination. In this spirit, they suggest several alternative publishing approaches and encourage the scientific community to explore and incorporate these into research assessment practices. They acknowledge, however, that metrics based on time and effort are inherently more complex and harder to interpret than those based on straightforward counts.

In conclusion, Trueblood et al. (2025) urge the research community to reshape academic publishing to better serve researchers and academia. Reforming the landscape of scientific publishing naturally leads to implications for research assessment.

References

DORA (2012). San Francisco declaration on research assessment. DORA – ASCB. Available from http://www.ascb.org/dora

Fyfe, A., Coate, K., Curry, S., Lawson, S., Moxham, N., & Røstvik, C.M. (2017). *Untangling academic publishing: A history of the relationship between commercial interests, academic prestige and the circulation of research*. Discussion Paper. University of St Andrews. https://eprints.bbk.ac.uk/id/eprint/19148/. Accessed April 11, 2025.

Gauffriau, M. (2021). Counting methods introduced into the bibliometric research literature 1970–2018: A review. *Quantitative Science Studies*, 2(3), 932–975.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429-431.

Larivière, V., Kiermer, V., MacCallum, C.J., McNutt, M., Patterson, M., Pulverer, B., Swaminathan, S., Taylor, S., & Curry, S. (2016). A simple proposal for the publication of journal citation distributions. bioRxiv preprint doi: https://doi.org/10.1101/062109

Perakakis, P., Taylor, M., Mazza, M., & Trachana, V. (2010). Natural selection of academic papers. *Scientometrics*, 85(2), 553–559.

Rousseau, R. (2015). The tip of the Chinese publication iceberg. *ISSI Newsletter* #44, 11(4), 100-102.

Seglen, P.O. (1989). From bad to worse: Evaluation by journal impact. *Trends in Biochemical Sciences*, 14(8), 326-327.

Seglen, P.O. (1997). Why the impact factor of journals should not be used for evaluating research. *BMJ*, 314(7079), 498-502.

Suiter, A.M., & Sarli, C.C. (2019). Selecting a journal for publication: Criteria to consider. *Missouri Medicine*, 116(6), 461-465.

Trueblood, J.S., Allison, D.B., Field, S.M., Fishbach, A., Gaillard, S.D., Gigerenzer, G., Holmes, W.H., Lewandowsky, S., Matzke, D., Murphy, S.D., Musslick, S., popov, V., Roskies, A.L., ter Schure, J., & Teodorescu, A.R. (2025). The misalignment of incentives in academic publishing and implications for journal reform. *Proceedings of the National Academy of Sciences of the United States of America*, 122(5), e2401231121.

Zhang, L., Rousseau, R., & Sivertsen, G. (2017). Science deserves to be judged by its contents, not by its wrapping: Revisiting Seglen's work on journal impact and research evaluation. *PLoS ONE*, 12(3): e0174205.