

Houston, we have a problem: Enhancing academic freedom and transparency in publishing through post-publication debate

Abstract:

This is submitted for the PSR symposium on academic freedom.

Debates over controversial articles often highlight important issues regarding academic freedom, transparency, and how to handle disagreement in publishing. I argue that a response outlining criticism is generally a more productive course of action than calling for retraction. However, there are a number of constraints that impede meaningful debates, and a problematic divergence between our common ideals of open research and free debate and the actual practices that we see in academic publishing, where our current practices often undermine transparency, replication, and scientific debate. I argue that research can benefit from more explicit recognition of politics and preferences in how we evaluate research as well greater opportunities for post-publication debate. The successful initiatives to promote data replicability over the past decade provide useful lessons for what improved post-publication transparency may look like.

Introduction

The article “The Case for Colonialism” originally published in *Third World Quarterly* in 2017 generated a controversy that highlights important issues regarding academic freedom, transparency, and how to handle disagreement in academic publishing. Some participants in the controversy simply called for the article to be retracted after publication, on the grounds that the article was offensive and damaging.¹ Retraction is normally an action reserved for articles that contain clear evidence of misconduct, usually understood as fraud or formal mistakes.² Calling for retraction for articles that one disagrees with is clearly problematic on grounds of academic freedom, commonly understood as the right of researchers to have freedom in conducting their research and seeking to publish the results.³ As many other participants in the ensuing debate, I consider the most appropriate response to articles you disagree with to write a response, outlining what you find problematic, and then allowing others to make up their mind about the issues at stake and the source of the disagreements.⁴ In this case, the original article is now officially withdrawn at the request of the author.⁵ As such, the article strictly speaking no longer exists in the public

¹ See <https://www.change.org/p/third-world-quarterly-call-for-apology-and-retraction-from-third-world-quarterly>.

² See, e.g., the guidelines by the Committee on Publication Ethics at https://publicationethics.org/files/retraction%20guidelines_0.pdf, and the discussion on RetractionWatch at <https://retractionwatch.com/2010/08/03/why-write-a-blog-about-retractions/>.

³ See, e.g., the American Association of University Professor’s “1940 Statement of Principles on Academic Freedom and Tenure”, available at <https://www.aaup.org/report/1940-statement-principles-academic-freedom-and-tenure>.

⁴ This view was expressed, among others, by Noam Chomsky, see <https://www.thecollegefix.com/post/36998/>.

⁵ For overviews of the calls for retraction and the petition of the author to have the article withdrawn, see for example <https://www.insidehighered.com/news/2017/09/26/author-third-world-quarterly-article->

domain and cannot be cited as a publication in the journal, although copies can still be found many places on the internet.⁶ It is certainly hard to publish a rebuttal to an article that formally no longer exists, and this makes it more difficult for outsiders to access the material, make up their own mind about the content, and to have an open public debate.

My main aim here is not to return to a debate on this specific article or event, but rather to point out how many issues of the issues raised in this controversy have broader relevance beyond this specific incident and indicate clear problems in our opportunities to have meaningful debates about research. Although I am very much in favor of the rebuttal over retraction approach to all research controversies, this is in practice often made much more difficult in practice than it ought to be. I will argue that there is a very problematic divergence between our common ideals of open research and free debate and the actual practices that we see in academic publishing, and that our current practices often undermine transparency, replication, and scientific debate.

When discussed in the abstract, there is a tendency to conceive of the research process in a rather idealized manner. Individuals make discoveries that they subsequently publish, and researchers can then debate the findings in a transparent manner. Hence, when researcher B disagrees with the prior work of researcher A, researcher B can try to replicate the original work of researcher A and publish a comment. In the end, both contributions will be evaluated in a transparent manner, purely based on their scientific merits, using criteria such as logical consistency and empirical content, free from any prior biases.

Of course, the real world differs from this idealized world in a number of respects. In the real world, it is often very difficult to find the original data from Researcher A in the first place. Moreover, there is often a great reluctance of editors to publish comments on the article by Researcher A submitted

colonialism-wants-it-stricken-record-it-might and <https://retractionwatch.com/2017/10/09/credible-threats-personal-violence-editor-prompt-withdrawal-colonialism-paper/#more-52049>.

⁶ See, e.g., http://www.web.pdx.edu/~gilleyb/2_The%20case%20for%20colonialism_at2Oct2017.pdf or <https://www.pdf-archive.com/2017/09/14/9-12-2017-the-case-f/>.

by Researcher B, and a substantial bias in favor of articles that already have been published and undergone peer review. Ideals are useful in outlining what we would like to see happen. But when the divergence from practices becomes so great that our ideals become illusions then we should recognize that we have a problem, and how this can be detrimental to both academic freedom and science. Instead of burying our heads in the sand, we need to step up our efforts and try to overcome these problems and better align practices with ideals. In my view, the answer lies in facilitating post-publication research and debate.

Illusion 1: Let us examine the evidence and publish a comment

My own experiences from research on polygyny and conflict provides a useful example of some of the problems often encountered in seeking to publishing comments on published research. In 2009, Kanazawa published an article in the *Journal of Politics* on what he called “the first law of intergroup conflict”. According to Kanazawa, contrary to common conceptions that see civil war as caused by enabling opportunities or political grievances, it is really polygyny leading to a scarcity of women that constitutes the ultimate cause of domestic conflict and violence. In defense of this thesis, Kanazawa claimed to provide empirical evidence showing that polygyny was consistently associated with civil war, while standard explanatory factors like income and democracy did not appear to be related to conflict in his analysis. Some colleagues and I noted a number of controversial design choices and possible problems in the analysis, and we eventually published as an article with a reply and some alternative analyses casting doubt on Kanazawa’s claims in the original journal [reference removed]. However, the process leading up to this illustrates important challenges in trying to publish rebuttals to published research.⁷

⁷ Kanazawa later gained prominence over a blog post where he discussed racial differences in a survey where subjects ranked “physical attractiveness”, see <http://www.bbc.co.uk/news/uk-england-london->

The *Journal of Politics* at the time had no replication requirement, and the data for the article were not publicly available. When we emailed Kanazawa about access to the data, we received a reply stating that “I am afraid that, due to a large number of attempts at theft, fraud, and other deceptive activities in the past, I can no longer share my data with anyone that I do not personally know. I'm sorry.” We included the original correspondence with Kanazawa in the first version of the article to explain why we were unable to replicate his analyses, but we were asked to take this out by the editors before they would agree to send the manuscript out for review. Although I am happy that we were eventually able to publish the article and comment on Kanazawa’s published claim, I still think that the email exchange was relevant information to reviewers and readers as to why we could not examine the original analysis, and I regret that it was not included in the comment.⁸

In many cases scientific debates become stifled as editors refuse to publish comments altogether. Many journals have a policy not to publish comments on articles at all, while other journals follow procedures that in practice make it very difficult to get comments that challenge published findings into

14945110. The public reaction led to a formal inquiry by his employer, the London School of Economics and Political Science (LSE). The LSE concluded that Kanazawa had brought the school into disrepute, and imposed sanctions, including a ban from teaching and publishing in non-peer-reviewed outlets for 12 months. I disagree with the conclusions and views expressed by Kanazawa, but submitted a letter to *The Guardian* at the time, defending his right to academic freedom and called for critics to provide a rebuttal rather than calling for sanctions. *The Guardian* did not publish the letter, but a copy is available at [reference removed].

⁸ For a copy of Kanazawa’s email, see [reference removed]. Following this incident, I and other members of the editorial board launched an initiative to lobby for a replication requirement for the *Journal of Politics*. This was eventually introduced, although the editors were not altogether enthusiastic about the idea, and expressed a concern that they did not want to be “data cops”.

print.⁹ In many cases scholars have discovered clear cases of mistakes, in various degrees of egregiousness and influence, after publication, including cases involving very prominent articles. For example, Reinhart and Rogoff (2010) published an article that claimed to find evidence that high debt to GDP ratios leads to reduced growth in a leading journal in Economics, and this was often cited as a rationale for austerity policies and the need to curb debt to promote future growth. Subsequent analyses have claimed that the conclusions do not follow from the evidence presented on a number of levels (for example, debt may be a result rather than a cause of low growth), but most notably the data analysis contained clear formal errors in the way the data had been handled in Microsoft Excel.¹⁰ However, no formal comment or debate has appeared on the article in the original journal outlet, even if at least one comment has appeared in another journal (see Herndon et al. 2014) and the authors have acknowledged the error elsewhere.¹¹ From a tabloid or “newsworthy” angle, it is easier to get comments noted or actually published when the original articles are prominent and have generated more attention. But from an open science point of view, the fact that an article may be less prominent or the errors less clear or egregious does not mean that it is less important to document challenges and criticism. In Economics the problem of constraints on debate of published articles is considered so severe that a separate journal *EconJournal Watch* has been set up to allow for debates on work published elsewhere.

⁹ For example, *Science* has followed a procedure for comments where the original author is first invited to write a response, before the original comment and the subsequent response is sent to an external reviewer. Although the original author often has important knowledge, she or he is not a disinterested participant, and the approach places a stronger bias against the publication of comments.

¹⁰ See <https://theconversation.com/the-reinhart-rogoff-error-or-how-not-to-excel-at-economics-13646> and <http://www.bbc.co.uk/news/magazine-22223190>.

¹¹ See <https://retractionwatch.com/2013/04/18/influential-reinhart-rogoff-economics-paper-suffers-database-error/>.

Ultimately, the career rewards for publishing comments are also limited. Comments are often not considered original research, and contributions will not be given as much weight as other published items. Indeed, some scholars would insist that any time spent on comments is detrimental for career prospects, especially for untenured junior scholars. In short, the sociology of the scientific practices we see in the real world differ dramatically from the textbook model of open science and scholarly debate, and in practice discourages rather than encourages efforts to replicate, reexamine, and reanalyze published findings.

Illusion 2: Debates and controversies are best settled by peer review

Other critics of the TWQ article have argued that the article should have been retracted due to alleged irregularities in the prior peer review process. More specifically, some argue that the article either was rejected by the reviewers, or “should have been”, and conclude that it therefore should not have been published. Peer review is a useful institution, and we are clearly much better off relying on peer review than the idiosyncratic judgement of a single editor. However, peer review is not a magical solution to the discovery of truth, and many things can go wrong during the peer review process from the perspective of academic freedom or open science.

Andrew Gelman has raised attention to the problem of “researcher degrees of freedom”. In brief, the apparent statistical significance of results presented alone cannot be a suitable measure of the strength of the evidence for a proposition, given the many opportunities for researchers to selectively change model specifications, collect additional data, or modify the initial research question and proposition being tested before the final analysis is presented.¹² However, in the peer review process we also have important forms of “editor degrees of freedom” and “reviewer degrees of freedom”. Editors have a great amount of freedom in how they solicit and use reviews. If editors are ill disposed to an article they can select

¹² See <http://andrewgelman.com/2012/11/01/researcher-degrees-of-freedom/>

reviewers that are more likely to be biased against it. Beyond intentional biases, editors may rely on reviewers that know little about the subject at hand or are unlikely to raise many objections. Articles may be deemed formally correct, but not considered “a sufficient contribution” to merit publication. Moreover, there is little to prevent reviewers from choosing their own criteria in deciding whether an analysis is “correct” or a “worthy” of publication in the first place. Most social scientists can probably cite examples where reviewers have recommended rejection primarily because they disagreed with finding on political grounds, or expressed that they were “uncomfortable” with how the findings in a manuscript could be used, without coming up with a principled reasons for why the results were incorrect or misleading. Hence, the cards are again often stacked in favor of the status quo, and stacked against open scientific debate on controversial topics.

The fact that something has passed peer review is no guarantee that the research must be correct. There is a growing awareness that the tabloid nature of publishing - where unusual and counterintuitive findings are often rewarded and preferred - can lead to publication biases where most published research findings are likely to be false (Ioannidis 2005). There is also a tendency to assume that peer review is based on a clear consensus view of quality and convergence in assessing research that simply does not exist. In a heterogeneous pool of reviewers the final replicability of decisions on acceptance and rejection can vary considerably depending on the draws of reviewers, making it less likely that peer review will identify what would otherwise be considered “the best” manuscripts (Esaray 2017). As in all binary tests, there are type I and type II errors in decisions on acceptance or rejection in peer-review and editorial decisions. Overall, many editors do a great job with considerable demands, despite limited resources and very few rewards for their efforts. However, we should not imbue the peer-review processes with magical powers for identifying truth that it cannot realistically meet. The fact that peer review is not infallible makes opportunities for post publication debate even more important.

A plea for transparency and post-publication discussion

If we agree that we have real problems with transparency and post-publication debate, how can we do better? I have no magic solution to offer here, but I think we can benefit from a much greater recognition of the role of politics and preferences in research as well as by efforts to expand opportunities for post-publication debate. Despite the problems highlighted here, I am fundamentally optimistic about research, and believe that we have made a great progress towards greater respect for academic freedom and research transparency. In particular, the new emphasis on replication requirements and data access has helped advance research transparency (see King et al. 1995; Gleditsch and Janz 2016). This idea was controversial at the time when first introduced, but much less so at the present, and I suspect that the benefits of replication criteria will be clear to most, including initial skeptics. This in turn has useful lessons for what increased post-publication transparency may look like.

Just as we now require researchers to show us the data, we should also demand to have more transparency about issues that do not follow from data and opportunities for debate. In some cases post-publication debate may revolve around what kind of data have not been used, or what the analyses may have looked like if certain assumptions are revised. Post-publication debate can also extend to policy implications and what kind of evidence can be the legitimate basis for policy. Many views about research, significance, and implications ultimately come down to differences in the initial preferences and views. Research can only be policy relevant given objectives, but the objectives themselves must be specified prior to the research process. Of course, research can often provide useful insights into means/ends relationships and the feasibility of attaining specific objectives. However, it is a fundamental mistake to present policy implications as “emerging” from research without specifying objectives in advance. It is unrealistic to expect that disputes over preferences and political views can be avoided in research altogether. In the real world, people’s initial political beliefs in practice have an important role in what topics they find interesting and acceptable topics for research as well as how they arrive at and consider research findings. There is nothing inherently wrong with this. It is perfectly acceptable to insist that certain research questions can be ethically problematic and for researchers to be politically engaged. However, we should recognize such concerns as political decisions and values and treat them as such,

rather than try to hide them behind appeals to science. We would all be much better off by being as explicit as possible in trying to debates about preferences from debates about empirical or logical issues and expanding the opportunities for post-publication discussion and providing recognition for efforts in this area. Although we have taken steps forwards on transparency when it comes to replication, we have perhaps taken steps backwards when it comes to allowing for comments. The challenges are perhaps understandable in an area of a dramatic increase in submission, limited journal space, acceptance rates in single digits, and difficulties in finding reviews. However, stifling postpublication debate is a recipe for disasters and undermining research transparency. We need both more outlets for postpublication debate and replication and recognize the value of engaging in such efforts.

References

Esarey, Justin. 2017. Does Peer Review Identify the Best Papers? A Simulation Study of Editors, Reviewers, and the Scientific Publication Process. *PS: Political Science & Politics* 50(4): 963-69.

Gleditsch, Nils Petter and Nicole Janz. 2016. Replication in International Relations. *International Studies Perspectives* 17(4): 361–66.

Herndon, Thomas, Michael Ash and Robert Pollin. 2014. Does High Public Debt Consistently Stifle Economic Growth? A Critique of Reinhart and Rogoff. *Cambridge Journal of Economics* 38(2): 257-79.

Ioannidis, John P. A. 2005. Why Most Published Research Findings Are False. *PLoS Med.* 2(8): e124, doi: 10.1371/journal.pmed.0020124.

Kanazawa, Satoshi. 2009. Evolutionary Psychological Foundations of Civil Wars. *Journal of Politics* 71(1):25-34.

King, Gary. 1995. Replication, Replication. *PS: Political Science and Politics* 28 (3): 444-452.

Reinhart, Carmen M. and Kenneth S. Rogoff. 2010. Growth in a Time of Debt. *American Economic Review* 100(2): 573-78.