A Descriptive Analysis of How Psychology Journals Handle Post-Publication Critiques

Whamond, A., Vazire, S., Clarke, B., Moodie, N., Schiavone, S., Thibault, R. T., & Hardwicke, T. E.

07 November, 2024

Abstract

Post-publication critique, such as letters-to-the-editor, can contribute to the validity and trustworthiness of scientific research. We conducted a cross-sectional and (b) prominent psychology journals. In 2023, an explicit submission option for post-publication critique was available at 23% (95% CI [16% to 32%]) of randomly sampled psychology journals (N = 100) and 38% of the most prominent psychology journals (N = 100). Journals sometimes imposed limits on the length and time allowed to submit critiques. We manually inspected two random samples of empirical articles published in 2020 (N = 101 articles per sample), estimating the prevalence of post-publication critique to be 0% (95% CI [0% to 3.7%]) in psychology journals generally and 1% (95% CI [0.2% to 5.4%]) in the most prominent psychology journals. The policy and practice of post-publication critique is seriously neglected in psychology journals.

Research transparency statement

Introduction

Critical discourse can contribute to the validity and trustworthiness of scientific research (Merton, 1973; Vazire & Holcombe, 2022). Criticism can highlight errors, limitations, and alternative interpretations, helping research consumers to calibrate their beliefs in scientific claims. However, the proliferation of poor-quality research in psychology (and other fields) indicates that the dominant mode of critical discourse, conventional peer review, is not sufficient to maintain high standards in the scientific literature (Hardwicke et al., 2020; Nosek et al., 2022).

Scientific journals can support the ongoing critical scrutiny of research by offering avenues for post-publication critique, such as letters to the editor (Douglas G. Altman, 2002). Journals' obligation to facilitate post-publication critique is explicitly codified in the principles of the Committee on Publication Ethics (COPE), an organisation that most prominent journals are members of (COPE, 2021; Hardwicke et al., 2022). However, prior research indicates that many (37%) of the most prominent academic journals across a broad range of scientific disciplines do not explicitly allow submissions of post-publication critique (Hardwicke et al., 2022). Moreover, journals that do allow post-publication critiques in principle rarely publish them in practice, and often impose restrictive limits on length and time-to-submit. The present research closely follows the methods used in Hardwicke et al. (2022) but extends this prior work by providing a focused evaluation of how psychology journals in particular handle post-publication critique (Hardwicke et al. examined 15 journals in a combined psychiatry and psychology category, along with 315 journals from other disciplines).

For the purposes of the present study, we operationally defined post-publication critique as any journal-based avenue for sharing peer-initiated critical discourse related to specific research articles previously published in the same journal (see Supplementary Information B for a complete operational definition). A prototypical example of post-publication critique is a letter to the editor, though there are other variations, such as commentary articles and website comments.

Our study was conducted in two parts. The first part aimed to describe policies related to post-publication critique at: (a) all psychology journals publishing empirical research (via a random sample of 100 journals), and (b) the 100 most prominent[1] psychology journals publishing empirical research. Specifically, we identified whether options for submitting critiques (e.g., letters to the editor) are available, and whether limits are imposed in terms of length, time-to-submit since publication of the target article, number of references, and whether critiques were sent for independent peer review. The second part of our study aimed to estimate the prevalence of post-publication critiques published in: (a) all psychology journals (via a random sample of 100 empirical psychology articles), and (b) the 100 most prominent psychology journals (via a random sample of 100 empirical articles published in these journals).

---

[1]There is no agreed upon or objective measure of journal prominence. To avoid an entirely subjective and arbitrary definition, we used Journal Impact Factor as a proxy for journal prominence. Journal Impact Factor, like all journal metrics, has limitations (especially when used as a proxy for quality, which is not what we are doing here), but using this method yields a collection of journals that we believe has good face validity in terms of prominence.

Part One: Post-Publication Critique Policies

Method

The methods and analysis plan were preregistered (https://osf.io/d6xf2). All deviations from the preregistered protocol were considered minor and are detailed in Supplementary Information A and, where relevant, mentioned below.

Design

This investigation had a cross-sectional design. All measured variables are described in detail in Supplementary Table C1. In brief, we extracted and described journal policies related to post-publication critique (according to our operational definition, Supplementary Information B) provided on journal websites in 2023. We recorded the name and description of any options for post-publication critique offered by each journal (e.g., article types such as 'letters' or 'commentaries', or web comments) or any explicit statement that the journal does not accept post-publication critique; limits imposed on post-publication critique in terms of length (e.g., number of words), time-to-submit (e.g., weeks since publication of the target article) and/ or number of references; and whether it was stated that post-publication critiques would be sent for independent external peer review (i.e., reviews solicited from individuals who were not members of the editorial team or authors of the target article). We recorded qualitative as well as quantitative limits; for example, a length limit might be stated as "500 words" or "short". We also identified whether journals were members of COPE, an organisation which outlines several core publishing practices, including the importance of facilitating post-publication critique (COPE, 2021).

Sample

Target populations

The target populations of the two samples of journals were: (a) all psychology journals publishing empirical research in 2023; and (b) prominent journals in psychology publishing empirical research in 2023. We were only interested in legitimate (non-predatory), peer-reviewed journals.

Sample sizes

The sample consisted of 100 randomly selected journals and 100 prominent journals. Justification for these sample sizes was based on a precision analysis (details provided in Supplementary Information E). Eleven journals appeared in both samples and were not replaced in order to maintain the random sampling scheme.

Preparing samples (inclusion/exclusion criteria)

We only included journals that were in the Web of Science Core Collection, meaning they met 24 basic quality criteria, such as having a peer-review policy and publishing scholarly content (for details see https://perma.cc/C5J3-V7C4). As of 15th March 2023, there were 867 unique journals included in this database that were classified as belonging to a subject category related to 'psychology'. Of these 867 journals, 115 were classified by Web of Science as non-English language and excluded, leaving 752 journals remaining. For the random sample of psychology journals, a list of all included 752 psychology journals was randomly shuffled using the R function 'sample' and we selected the first 600 on the shuffled list (to allow for exclusions). For the prominent journals sample, we sorted journals in order of their 2021 Clarivate Journal Impact Factors (https://jcr.clarivate.com/) and selected the first 600 journals (to allow for exclusions). The list of these journals is available at https://osf.io/k6yx5. For both samples, journals were then evaluated sequentially from the top of their respective list until the target of 100 eligible journals in each sample was reached. If a journal needed to be excluded after manual inspection, it was replaced with the next available journal in its list, thus maintaining the respective sampling schemes. We excluded an additional 6 journals from the randomly-selected journal list and 15 from the prominent list because they did not publish empirical research. This determination involved some degree of subjective judgement and was primarily based on each journal's scope as defined on its website. If the journal's description of its scope was ambiguous, we searched the journal's archive for examples of empirical articles to inform our determination. This represented a slight deviation from the preregistered protocol (for details see Supplementary Information A).

Procedure

Piloting and training

The data extraction procedures were piloted with 10 journals that were not included in the sample. All investigators were trained and familiarised with the data extraction process and practised with positive and negative training examples.

Collating journal policy information

One investigator (AW or NM) identified, downloaded, and archived the 'scope' and 'article types' sections (or equivalent) provided on each journal's website (https://osf.io/8k7m4/). This information was used to determine whether the journal published empirical articles and, if so, whether post-publication critique article types were accepted by the journal. If these sections could not be found, a second investigator (TEH) double-checked.

Data extraction

The policy information (outlined in Supplementary Information C using a Goole Form (https://osf.io/d6uzj) was extracted and classified for each journal by two investigators. Primary coding was conducted by AW (n = 183) and NM (n = 37). Secondary coding was conducted by RTT (n = 82), SV (n = 80), and TEH (n = 41). Disagreements between primary and secondary coders, usually arising due to unclear descriptions of article types, were resolved by TEH. This was necessary for 40 journals (for details see Supplementary Information H). For coding whether journals provided an option for post-publication critique, we first examined the journal website's description of the journal's article types (typically found in journals' submission guidelines or instructions to authors). If a journal's description of its article types was not provided or unclear, we searched the journal's archive to see if any examples of articles meeting our operational definition of post-publication critique (Supplementary Information B) had been published within the last five years (i.e., since the beginning of 2019). Note that if a journal had an article type that sounded like it might be used for post-publication critique (e.g., "Commentary"), that was not sufficient for a positive classification; the description of the article type, or examples of that article type, also had to meet our operational definition of post-publication critique (see Supplementary Information B).

Data harmonization

Journals used various naming conventions for similar types of post-publication critique and various units (e.g., characters, words or pages for length limits) to specify limits. To address this, we harmonised the various names used for critique formats to four types (letters [to the editor], commentaries, web comments, and other) and transformed length limit units to words and time-to-submit limit units to weeks (for details see Supplementary Information G).

Results

Journal characteristics

The median 2021 Journal Impact Factor was (interquartile range = 1.83 - 3.81) for the randomly sampled journals (NB. 25 of the randomly selected journals did not have Journal Impact Factors) and 6.92 (interquartile range = 5.6 - 8.69) for the prominent journals. 65 of the 100 randomly sampled journals, and 84 of the 100 prominent journals, were members of COPE. Journal characteristics stratified by psychology sub-field are available in Supplementary Information I.

How many journals offer post-publication critique?

A full list of journals and their post-publication critique options and limits is provided in Supplementary Information J (raw data available at https://osf.io/8k7m4/). Among the 100 randomly-selected journals, 23 (23%, 95% CI [16%, 32%]) explicitly stated that they allowed submissions of post-publication critique. Of these 23 journals, had two types of critique formats and the other 21 each had one critique format. journal (*Clinical Psychology: Science and Practice*, also included in the prominent journals sample) stated explicitly that they did not accept post-publication critique. 12 journals offered commentaries, 12 journals offered

letters, 0 journals offered web comments, and journal (*Cortex*) offered a format called 'verification reports' which we classified as 'other'. Of the 23 journals that explicitly accepted post-publication critique, 15 (65%) were COPE members. Of the 50 journals that did not have a policy related to post-publication critique, 50 (65%) were COPE members.

Among the 100 prominent journals, 38 (38%) explicitly stated that they allowed submissions of post-publication critique[2]. Of these 38 journals, had two types of critique format and the other 35 journals had one critique format. journals (*Clinical Psychology: Science and Practice* and *Psychological Bulletin*) stated explicitly that they did not accept post-publication critiques. No reason was provided. 26 journals offered commentaries, 16 journals offered letters, and 0 journals offered web comments. Of the 38 journals that explicitly accepted post-publication critique, 33 (87%) were COPE members. Of the 51 journals that did not have a policy related to post-publication critique, 51 (82%) were COPE members.

What limits did journals place on post-publication critique?

Table 1 (randomly selected journals) and Table 2 (prominent journals) show the limits journals imposed on post-publication critique in terms of length, time-to-submit, and number of references. The tables also show whether a peer review policy for critiques was stated. Length limits were stated for most critique formats, especially those in prominent journals. Length limits were usually stated quantitatively (e.g., "1,000 words"). A few formats had qualitative limits (e.g., "brief"). Letters had tighter length limits than commentaries on average. The most restrictive length limit was 400 words at the *Journal of Adolescent Health*. Time-to-submit limits were not stated for most critique formats; they were somewhat more common in prominent journals. Time-to-submit limits, when present, were often stated qualitatively (e.g., "brief") and sometimes quantitatively (e.g., "3 months"). The most restrictive time-to-submit limit was 4 weeks at the *Journal of Applied Psychology*. For most critique formats, journals did not state whether they would be sent for independent peer review or not. For a full list of stated limits at each journal, see Supplementary Information J.

---

[2]Confidence intervals are not included for the prominent psychology journals as this sample represents our entire population of interest, negating the need for inferential statistical calculations.

Table 1. Post-publication critique types identified in randomly selected psychology journals and their length, time-to-submit, and reference limits.

Post-publication critique type	Length limits			Time limits			Reference limits			Peer review
	any n (%)	quantitative n (%)	Median (IQR)	any n (%)	quantitative n (%)	Median (IQR)	any n (%)	quantitative n (%)	Median (IQR)	'Yes' n (%)
Commentary	9 (75)	9 (75)	2000 (1200 - 2000)	4 (33)	1 (8)	39 (39 - 39)	4 (33)	4 (33)	12.5 (9 - 15)	2 (17)
Letters	3 (25)	3 (25)	1000 (750 - 1750)	0 (0)	0 (0)	NA (NA - NA)	0 (0)	0 (0)	NA (NA - NA)	0 (0)
Other	0 (0)	0 (0)	NA (NA - NA)	0 (0)	0 (0)	NA (NA - NA)	0 (0)	0 (0)	NA (NA - NA)	0 (0)
All types	12 (48)	12 (48)	1600 (1075 - 2125)	4 (16)	1 (4)	39 (39 - 39)	4 (16)	4 (16)	12.5 (9 - 15)	2 (8)

Note. This table also indicates whether the journal states that critiques will be sent for independent external peer review. IQR = Interquartile Range. Note that n's refer to number of critique formats, not number of journals (some journals had more than one format of post publication critique). All results rounded to the nearest integer.

Table 2. Post-publication critique types identified in prominent psychology journals and their length, time-to-submit, and reference limits.

Post-publication critique type	Length limits			Time limits			Reference limits			Peer review
	any n (%)	quantitative n (%)	Median (IQR)	any n (%)	quantitative n (%)	Median (IQR)	any n (%)	quantitative n (%)	Median (IQR)	'Yes' n (%)
Commentary	21 (81)	18 (69)	2000 (1200 - 2912)	8 (31)	4 (15)	26 (11 - 39)	6 (23)	6 (23)	15 (11 - 15)	10 (38)
Letters	13 (81)	13 (81)	800 (500 - 1000)	4 (25)	2 (12)	39 (33 - 46)	9 (56)	9 (56)	10 (5 - 10)	3 (19)
All types	34 (81)	31 (74)	1200 (1000 - 2250)	12 (29)	6 (14)	33 (16 - 39)	15 (36)	15 (36)	10 (8 - 15)	13 (31)

Note. This table also indicates whether the journal states that the critiques will be sent for independent external peer review. IQR = Interquartile Range. Note that n's refer to number of critique formats, not number of journals (some journals had more than one format of post publication critique). All results rounded to the nearest integer.

Part Two: Post-Publication Critique Prevalence

Method

The methods and analysis plan were preregistered (https://osf.io/d6xf2). All deviations from the preregistered protocol were considered minor and are detailed in Supplementary Information A and, where relevant, mentioned below.

Design

This investigation had an cross-sectional design. All measured variables are described in detail in Supplementary Information C2. In brief, we estimated the prevalence of post-publication critique among empirical articles published in 2020 using two methods: (a) we identified whether sampled articles contained a link to a post-publication critique of the article ("Linked Method"); and (b) we identified whether sampled articles were themselves examples of post-publication critique ("Instance Method")[3]. These two methods of estimating prevalence have complementary strengths and weaknesses; for more details, see Supplementary Information D.

Sample

Target populations

The target populations were (a) all empirical articles published in 2020 in all psychology journals; and (b) all empirical articles published in 2020 in the 100 most prominent psychology journals identified in Part One. We were only interested in articles published in legitimate (non-predatory), peer-reviewed journals.

Sample sizes

We examined 101 randomly selected empirical articles published in psychology journals and 101 randomly selected empirical articles published in prominent psychology journals (our target sample size was 100 articles in each sample, but, deviating from our preregistration, we accidently examined one additional article in each sample, see Supplementary Information A for details). Justification for these sample sizes is provided in Supplementary Information E.

Preparing Samples (inclusion/exclusion criteria)

For the random sample of articles published in psychology journals, we used the Web of Science Core Collection to download bibliographic classified with a Web of Science document type of 'Article', 'Discussion', 'Editorial material', or 'Letter' published in 2020 and classified as belonging to a psychology research area. The search returned 61,735 records (for exact search string, see Supplementary Information F).

For the random sample of articles published in the most prominent psychology journals, we used the Web of Science Core Collection to download bibliographic records classified with a Web of Science document type of 'Article', 'Discussion', 'Editorial material', or 'Letter' published in 2020 in the 100 most prominent psychology journals (identified in Part One). The search returned 12,255 records (for exact search string, see Supplementary Information F).

The two lists of sampled articles were each randomly shuffled using the R function 'sample' and the top 600 entries in each shuffled list were selected (to allow for exclusions). Articles were then evaluated sequentially until the target of 100 eligible articles in each sample was reached. Articles were excluded because they: (i) could not be accessed (n = 9); (ii) were non-English language (n = 4); (iii) indicated they had been retracted (n = 2), (iv) were non-empirical (n = 55), or (v) were conference abstracts or 'supplement' articles rather than complete articles (n = 9). Assessing whether an article is empirical involved some subjective judgement, but we used the following guidelines: empirical research must include some analysis of quantitative or qualitative data and the following do not count as empirical research: news, book reviews, editorials, previews, opinion

---

[3]In Hardwicke et al. (2022), the Linked Method and Instance Method of prevalence estimation were called the 'primary' and 'secondary' method respectively. We changed the terminology because we do not want to imply that one method is necessarily superior to the other.

pieces, simulations, proofs, theoretical papers, and reviews. We included primary research, secondary data analysis, and meta-analyses. If an article needed to be excluded after manual inspection, it was replaced with the next available article in its list, thus maintaining the random sampling scheme.

Note, the above applies to the Linked Method for calculating prevalence; for the Instance Method for calculating prevalence estimates, articles that were themselves post-publication critique (e.g. 'Letters') were included, but they did not count towards the 100 empirical articles sample size target.

Procedure

Piloting and training

The data extraction procedures were piloted with 10 articles that were not included in the sample. All investigators were trained and familiarised with the data extraction process and practised with positive and negative training examples.

Data extraction

To code whether sampled articles were linked to post-publication critique (Linked Method), two investigators independently viewed each article on the journal's website, assessed if the article met the eligibility criteria, and if it did, thoroughly checked the page for any linked post-publication critique.

To code whether the sampled article itself was an example of post-publication critique (Instance Method), the same two coders also read the title and, if necessary, the abstract of the sampled article.

Primary coding was conducted by AW (n = 246) and NM (n = 35). Secondary coding was conducted by BC (n = 122), SS (n = 110), and TEH (n = 49). Data extraction/classification disagreements were resolved by TEH, which was necessary for 15 articles (for details see Supplementary Information H).

Results

Among the 101 empirical articles randomly sampled from all psychology journals, we did not find any examples of post-publication critiques based on the Linked Method nor the Instance Method, resulting in a prevalence estimate of 0% (95% CI [0%, 3.7%]) for both methods.

Among the 101 empirical articles randomly sampled from prominent psychology journals, we found only one post-publication critique using the Linked Method (i.e., one sampled article was associated with a linked post-publication critique), yielding a prevalence estimate of 1% (95% CI [0.2%, `rci_promLinked_UL%`]). `Using the Instance Method, we found no post-publication critiques (r prominent_instance%`, 95% CI [0%, 3.7%]) (i.e., none of the sampled articles were themselves instances of post-publication critique).

General Discussion

Post-publication critique is often viewed as an important mechanism of scientific self-correction (Merton, 1973; Vazire & Holcombe, 2022). However, our study found that across the field of psychology only a minority of journals (23, 23%, 95% CI [16%, 32%]) had an explicit submission format for post-publication critique, and actual examples of post-publication critique are extremely rare. Prominent psychology journals were more likely to have an explicit format for post-publication critique according to their policies (38%), but the majority still did not. No journals in either sample supported web comments. Journals that did allow submission of post-publication critique sometimes imposed restrictions on length and time-to-submit. Occasionally, these limits were quite restrictive; the strictest length limit was 400 words and the strictest time-to-submit limit was 4 weeks after the target article was published. In practice, post-publication critiques were rare with only a single Linked example found within the articles sampled from prominent journals. Overall, our findings suggest that the policy and practice of post-publication critique is seriously neglected in psychology journals.

Should all journals offer an explicit avenue for post-publication critique? Post-publication critique is recognized as an important mechanism of scientific self-correction (Merton, 1973; Vazire & Holcombe, 2022)

and facilitating critique is often considered to be a journal's responsibility (e.g., according to COPE (2021)). Interestingly, in both the randomly selected and prominent samples, most (65% and 82%, respectively) of the journals that did not offer post-publication critique were members of COPE[4]. It is unclear if publishers or journal editors have principled reasons for not explicitly offering post-publication critique.

It could be argued that post-publication critique can take place in other contexts, such as social media or dedicated platforms like PubPeer (Bastian, 2014); however, journals are in a privileged position to encourage, curate, and publish critiques in the scientific record, which improves permanence and discoverability (i.e., by linking to critiques on the same webpage as the target article). Critiques posted informally on external platforms may be overlooked by journals and readers (Allison et al., 2016). There is also a risk that third party sites will become defunct — as in the case of PubMed Commons, which was shut down in 2018 [ncbiinsights2018]. Some prominent psychologists have even argued that critiques should appear only in venues subject to editorial oversight (e.g., Fiske (2016)).

Are journal limits on post-publication critique justifiable? Length limits may promote concision; however, strict limits (e.g., a few hundred words) are likely to seriously reduce the scope and substance of critiques (Douglas G. Altman, 2005). There is no obvious justification for imposing time-to-submit limits. Such limits imply that there is effectively a "statute of limitations" on scientific critique once a paper has been published for a given time (often no more than a few months). Only a few policies stated that critiques would be sent for independent external peer-review. It was unclear whether journals routinely send critiques to the original authors for review, but anecdotally, we know this can occur. Independent peer-review can provide more impartial assessment of critiques and bolster their legitimacy Schriger & Altman (2010). Without independent review, the decision about whether to publish a critique may be based solely on the perspectives of individuals with a conflict of interest, namely the journal editors and the original authors.

Though our data indicate that post-publication critique is neglected in psychology journals, they do not provide insight into why this is the case. Potential reasons could be cultural (e.g., disciplines have different attitudes towards scientific criticism and how to handle it), pragmatic (e.g., some disciplines have a greater need for critique than others because of lower quality standards), bureaucratic (e.g., availability of resources to support post-publication critique), historic (e.g., particular individuals or events that have highlighted the value of post-publication critique), or reputational (e.g., a perception that post-publication critique could damage journal reputation). Other factors which may deter submission of post-publication critique include fear of retaliation, such as public defamation and legal threats from those whose research they questioned (Besançon et al., 2021).

Our research has several limitations. First, our operational definition of post-publication critique was intentionally focused on journal accountability and therefore some types of critical discourse (such as externally hosted critiques or criticisms embedded in research or review articles that do not target a specific article) were beyond the scope of this study. Our definition also excluded errata, corrections, and retractions, some of which may have been motivated by post-publication critiques that were eventually not published. Second, we relied on information explicitly published on journal websites; it is possible that some journals accept post-publication critique, even though they do not have an explicit policy saying so on their website. Third, it is possible that we missed some examples of post-publication critique because they were not clearly linked to the target article; we did our best to minimize this possibility by having two coders independently assess the presence of critique for each article. Fourth, our study focused on constraints imposed by journals on post-publication critique; however, there are likely many contributing factors to the low prevalence of critique, such as authors' ability and willingness to write and submit critiques, and the extent to which critiques are recognized in hiring, promotion, and funding decisions.

Conclusion

Since 2011, psychology has been navigating a period of considerable epistemological turmoil, with serious concerns raised about research quality (Nosek et al., 2022). Post-publication critique is a valuable mechanism

---

[4]COPE guidelines explicitly recommend that journals allow post-publication critique "on their site, through letters to the editor, or on external moderated website such as PubPeer" (COPE Council, 2021). Technically, a journal does not need to do anything to comply with this recommendation because PubPeer is an independent platform.

for identifying problems in published research and highlighting them to readers (Hardwicke et al., 2022; Merton, 1973; Vazire & Holcombe, 2022). The present study shows that post-publication critique is seriously neglected in policy and practice at psychology journals. Most journals do not have explicit formats for the submission of post-publication critiques and critiques are very rarely published. In Box 1, we provide some tentative policy suggestions that may help to improve this situation.

INSERT BOX 1

Author contributions

Conceptualization: T.E.H., A.W., S.V. Data curation: A.W. and T.E.H. Formal analysis: A.W. and T.E.H. Investigation: A.W., B.C., N.M., S.S., S.V, R.T.T., and T.E.H. Methodology: T.E.H. Project administration: A.W. and T.E.H. Resources: T.E.H. Software: T.E.H. Supervision: T.E.H. and S.V. Validation: A.W. and T.E.H. Visualization: A.W. Writing - original draft: A.W. Writing - review & editing: A.W., S.V., and T.E.H.

References

Allison, D. B., Brown, A. W., George, B. J., & Kaiser, K. A. (2016). Reproducibility: A tragedy of errors. *Nature*, *530*(7588), 27–29. https://doi.org/10.1038/530027a

Altman, Douglas G. (2002). Poor-Quality Medical Research: What Can Journals Do? *JAMA*, *287*(21), 2765. https://doi.org/10.1001/jama.287.21.2765

Altman, Douglas G. (2005). Unjustified Restrictions on Letters to the Editor. *PLoS Medicine*, *2*(5), e126. https://doi.org/10.1371/journal.pmed.0020126

Bastian, H. (2014). A Stronger Post-Publication Culture Is Needed for Better Science. *PLoS Medicine*, *11*(12), e1001772. https://doi.org/10.1371/journal.pmed.1001772

Besançon, L., Peiffer-Smadja, N., Segalas, C., Jiang, H., Masuzzo, P., Smout, C., Billy, E., Deforet, M., & Leyrat, C. (2021). Open science saves lives: Lessons from the COVID-19 pandemic. *BMC Medical Research Methodology*, *21*(1), 117. https://doi.org/10.1186/s12874-021-01304-y

COPE. (2021). *Handling of post-publication critiques*. Committee on Publication Ethics. https://doi.org/10.24318/o1VgCAih

Fiske, A. P. P. S. T. (2016). A Call to Change Science's Culture of Shaming. *APS Observer*, *29*.

Hardwicke, T. E., Serghiou, S., Janiaud, P., Danchev, V., Crüwell, S., Goodman, S. N., & Ioannidis, J. P. A. (2020). Calibrating the Scientific Ecosystem Through Meta-Research. *Annual Review of Statistics and Its Application*, *7*(1), 11–37. https://doi.org/10.1146/annurev-statistics-031219-041104

Hardwicke, Thibault, Robert T., Kosie, J. E., Tzavella, L., Bendixen, T., Handcock, S. A., Köneke, V. E., & Ioannidis, J. P. A. (2022). Post-publication critique at top-ranked journals across scientific disciplines: A cross-sectional assessment of policies and practice. *Royal Society Open Science*, *9*(8), 220139. https://doi.org/10.1098/rsos.220139

Merton, R. K. (1973). The normative structure of science. In *The Sociology of Science: Theoretical and Empirical Investigations.* (pp. 267–278). University of Chicago Press.

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., Fidler, F., Hilgard, J., Kline Struhl, M., Nuijten, M. B., Rohrer, J. M., Romero, F., Scheel, A. M., Scherer, L. D., Schönbrodt, F. D., & Vazire, S. (2022). Replicability, Robustness, and Reproducibility in Psychological Science. *Annual Review of Psychology*, *73*(1), 719–748. https://doi.org/10.1146/annurev-psych-020821-114157

Schriger, D. L., & Altman, D. G. (2010). Inadequate post-publication review of medical research. *BMJ*, *341*(aug11 3), c3803–c3803. https://doi.org/10.1136/bmj.c3803

Vazire, S., & Holcombe, A. O. (2022). Where Are the Self-Correcting Mechanisms in Science? *Review of General Psychology*, *26*(2), 212–223. https://doi.org/10.1177/10892680211033912