

Strasbourg, 5 February 2020

Dear Daniel,

Submission of “Reviewers’ Decision to Sign Reviews is Related to Their Recommendation” by
Nino van Sambeek & Daniël Lakens—*Meta-Psychology* manuscript number 2289

First, let me apologise for the length of time that it has taken to get back to you about this submission. Due to a mix-up between me and the editor-in-chief, I didn’t realise that I was being requested to edit the manuscript, as opposed to simply reviewing it!

I have received two reviews, from Sam Westwood of King’s College London and Cody Christopherson of Southern Oregon University. Both are broadly positive about your manuscript and consider the subject worthy of publication. They have made a number of individual points, which are detailed in the rest of this letter. Some of these are direct suggestions for revisions, while others engage with your manuscript in a more conversational way. Although there are quite a number of points, I don’t think that this meets the threshold for “major revisions”, so my formal decision is “minor revisions required”.

I have also included some notes of my own, which I initially made when I thought I was going to be a reviewer. At the same time I also made a few annotations to the PDF of the manuscript; I have uploaded this file to the “Review round 1” file area of the project’s OSF repository (<https://osf.io/w2t7u/>). As far as I have been able to determine, there are no notes in *hypothes.is* on your preprint.

Please read the reviewers’ comments carefully and for each one either revise your manuscript accordingly, or indicate why you don’t think this is necessary.

I look forward to receiving your reply and revised manuscript. There is no particular deadline, but if you anticipate taking more than around 60 days, please let me know.

Best regards,

Nick Brown (Guest editor)

Reviewer #1 (Sam Westwood)

Recommendation: Revisions Required

Reviewer Comments

(Text copied and pasted from the submission site by Nick Brown; I cleaned up the consequent automatic formatting a little, hopefully without introducing errors)

This is a timely investigation into the potential (unintended) consequence of open signed peer review on peer-reviewer's behaviour. The pattern of data shows that, when compared against non-signed reviews, signed reviews are associated with more positive recommendations.

Whether this is a causal association remains to be seen, but research in this area would suggest this indicates open-reviewers are demonstrating their commitment to openness while avoiding backlash from their colleagues. The authors recognise the investigation is exploratory and results cannot lead to any substantive causal inferences, which is made clear in a conservative and well-balanced argument.

After reading the article several times over, I kept asking whether the authors could have done more. At present, we have broad strokes, but guidance in future research could be helped if a more fine-grained summary of the data were given. Therefore, recommendations in my review are to dig deeper to unpack potentially useful information. I have asterisked comments which I deem to be major revisions.

***Comment #1**

I consider the following statement to be too strong: Our data support the idea that researchers decision to sign is related to their recommendations across a wide range of scientific disciplines. To justify this, I would like to see a breakdown of their findings by discipline, which appears to

be possible given the volume of articles and the interdisciplinary nature of the journals from which the articles were extracted. Regardless of the presence or absence of this statement, this is a worthwhile topic for further exploration and/or discussion. Disciplines differ in their peer review approach; some weigh methodological quality over a defensible theory/ inference. One might expect higher numbers of signed reviews for the former since the review is focused on reviewing a study on more objective criteria, thus any backlash might run the risk of appearing too personal if the methods are clearly flawed. Alternatively, some disciplines are more competitive than others, and where competition is high so too is the risk of backlash (or skulduggery on the part of the reviewer).

Comment #2

Have certain disciplines taken to sign peer review more so than others? Providing data on this would be informative for future researchers designing studies investigating the barriers/weaknesses to open peer-review. Addressing comment #1 would address this comment too.

*Comment #3

Did the authors explore reviewers that have given more than one review for PeerJ or Royal Society Open Science? One would assume that these reviewers would always sign their reviews, but what if this is not the case. If so, why did they review papers some of the time but not others? Was it because the one they signed was more positive? This is worth a mention. Unpacking this might show how systematic the association between the positivity of a recommendation is and whether the review is signed or not. This should at least be discussed as a limitation if the data is not available.

Comment #4

The identified articles were published over several years. Surely the incidence of signed reviews has changed over time? People might have been hesitant/unaware at the time of their introduction, but persuaded by how normative/acceptable open criticism -- or its discussion -- has become (of course this will be dependent on the discipline and level of seniority of the persons involved). If signed reviews have increased over time (as one might expect), it would add urgency to further investigation into policies incentivising such behaviour change, given their unintended (potential) effect.

*Comment #5

Journals sometimes give the authors of a submitted manuscript the option to suggest reviewers. If authors can nominate negative reviewers, one might suspect that the rate of positive signed reviews would be an underestimate. The authors should report on whether PeerJ or Royal Society Open Science have this policy, and what it might mean for policies of open peer review.

Comment #6

There is unlikely to be one reason why reviews go unsigned, and I think the authors pick the one most generalisable. However, backlash can come in different forms. A career-limiting critical peer review is one form of backlash, but another is revealing how peer review can be gamed. I work in developmental psychiatry, and a number of journals give the option of positive/negative reviews. Mandatory signed peer review would expose how authors target colleagues most likely to give favourable reviews. I invite the authors to reference this point in their article, but this is optional.

***Comment #7**

Please ensure the manuscript is proofread for clerical/grammatical errors. The word "individual" is misspelt, "researchers decision" is missing an apostrophe.

***Comment #8**

Please check the link to the following reference, it appears to be broken in the pdf I reviewed:
Bastian, H. (2018). Signing Critical Peer Reviews & the Fear of Retaliation: What Should We Do? |Absolutely Maybe. <https://blogs.plos.org/absolutely-maybe/2018/03/22/signing-critical-peer-reviews-the-fear-of-retaliation-what-should-we-do/>.

(Note by Nick: The points raised in comments #7 and #8 are flagged in the annotated PDF file.)

Reviewer #2 (Cody Christopherson)

Recommendation: Revisions Required

Reviewer Comments

(Text copied and pasted from the submission site by Nick Brown; I cleaned up the consequent automatic formatting a little, hopefully without introducing errors)

Thank you for the invitation to review “Reviewers’ Decision to Sign Reviews is Related to Their Recommendation”. In this paper, the authors seek to demonstrate a pattern in whether pre-publication reviews of scientific manuscripts are signed by the reviewers. To this end, they have scrapped and coded data from 16,198 open reviews. I have set up the remainder of this review in two sections—first, strengths of the article; second, suggestions for improvement. I consider these suggestions minor.

Strengths of the article:

- The authors take a novel approach (scrapping open reviews) to address a familiar question. This conversation has been in the literature at least 20 years but this appears to be the first large-scale attempt get descriptive statistics. The approach is ambitious and fruitful.
- Appropriately modest about the relationship between signature and positivity of the review, noting several important limitations (correlational and lacking an important part of the distribution, namely, rejected articles)

Suggestions for improvement:

*16,000+ articles is remarkable in and of itself. Because all the articles come from two journal families, it would be helpful to add a one paragraph description of PeerJ and RSOS. It might

include their scope, reputation, or any other information that would allow the reader to contextualize the sample.

*Add information to make the figures more distinct. Figure 2 can include the words “Royal Society”. This will make clear using the figures alone that figures 1 and 2 represent entirely separate samples.

*Under the subheading “Additional Analyses”, the authors offer some analysis regarding “how often reviewers agree.” This is an interesting approach to a common-place but under-studied phenomena in meta-science. That said, these analyses don’t contribute to the main theme of the paper, are not referenced in any other part of the paper, are not situated in the reviewed literature, and are not interpreted in any way. I suggest that this section be either removed or expanded and incorporated into the larger theme.

*Another angle to consider: the authors focus on signed positive reviews being incentivized (the reviewer gets credit without backlash from the authors) and signed negative reviews disincentivized (the reviewer get credit but may get backlash from the authors). However, with increased in retractions (Steen, Casadevall, & Fang, 2013) and post-publication criticism (via social media, PubPeer, and so on), I think that the risk in a signed positive review is understated in the article. Imagine if the public knew who reviewed the many high-profile retractions over the last few years because they could read their review for themselves. Reputations could easily be damaged by this type of transparency.

*The authors acknowledge that the causal relationship for the relationship they observe could run two different ways. Given the (well-acknowledged) incomplete nature of the data, considering at least one more counter-narrative would make for a more complete treatment of the issue.

*Not part of the article, but the cover letter does not make sense. I suspect it has an errant paragraph from a cover letter for a different article.

Editor comments (Nick Brown)

* It seems to me that there may be a possible source of bias in the decision by the author(s) of the published articles whether or not to have the reviews published alongside their article, which (according to the manuscript) is optional at *PeerJ* and was optional for a time at the Royal Society journals. Maybe this depends on an interaction between the identity of the reviewer and their recommendation (e.g., if a professional rival recommended rejection and the editor overruled them).

* There are no quantitative analyses in the manuscript other than proportions. Why is this? Maybe the authors think their findings are too exploratory to justify the presentation of p values, but some indication of the magnitude of the effects would surely be useful.

* I would be interested to see a short discussion of the possible consequences of the fact that at *PeerJ*, readers do not get to see the individual reviewers' recommendations.

* “Our data support the idea that researchers [*sic*] decision to sign is related to their recommendation across a wide range of scientific disciplines” (p. 4)—this language could suggest that a robust effect has been shown for each of several disciplines. However, all that has actually been shown (I think) is that there is some effect (cf. my earlier remark about effect sizes) across a sample that encompasses a wide range of disciplines. It could be that the effect is stronger for psychology and weaker or absent for biomedical sciences, for example.

* “For positive recommendations, reviewers will get credit for their reviews, while for negative reviews they do not run the risk of receiving any backlash from colleagues in their field” (p. 4)—I’m not sure if “credit” is the right word. Reviewers who write a glowing report on

a manuscript can certainly expect to be liked by its author(s), but whether or not the rest of their colleagues (and the field as a whole) will give them any sort of “credit” will depend on whether the article later turns out to be flawed (and/or whether it upsets some powerful people); and the same could apply in reverse for a negative review. So perhaps this comment should be rephrased to emphasise that any “credit” is restricted to approval felt from the author(s) of the accepted article.

* I downloaded the code and data associated with the manuscript via the OSF link and was able to reproduce the main numerical results and figures. However, the percentages in the first paragraph on page 4 were printed with a lot more decimal places than in the PDF file (e.g., “For 41.7218543% of the manuscripts the maximum deviation was one category”). I wonder if your R setup has a different global rounding option from mine? For best reproducibility, rounding should probably be done explicitly in the code.