

# Reply to reviews

Daniel Lakens

5/24/2020

## Reviewer 1

Comment #1 and #2 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. 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.

**Reviewer 1 notes how our statement that we found a pattern across ‘across a wide range of scientific disciplines’ was too strong. Although we simply meant to imply we found it when averaging over fields, we see the benefit of an analysis per subject area.**

We have added a plot the provides an overview of the data for different sections for PeerJ. Royal Society Open Science has different subjects that each have an editorial board (<https://royalsocietypublishing.org/rsos/editorial-board>) and distinguishes “collections” as can be seen at <https://royalsocietypublishing.org/rsos/collection>. However, neither of these classifications are part of the meta-data in scopus, nor are they in the reviews. Therefore, for Royal Society Open journals we can only split the analysis for Royal Society Open Science and Royal Society Open Biology, the latter of which has very few datapoints. The pattern is the same, but with 300 reviews, mostly unsigned, this analysis is not informative enough to be included in the manuscript (the code to generate the plot for Open Science and Open Biology is in the manuscript code, but commented out). The new figure that plots the data across sections in PeerJ also addresses the second comment by Reviewer 1, concerning whether some fields sign more than others. We note how in paleontology, signing is most frequent. The revised manuscript now reads, just above where we present the new figure plotting the result for each section:

*In general, reviews are less likely to be signed than unsigned, except for the paleontology and evolutionary science section, where reviews are more likely to be signed than unsigned. Beyond this main effect, we see a clear interaction across all sections, where the probability that a ratio of signed versus unsigned reviews is greater for major revisions than for minor revisions. Across all fields covered by PeerJ, it is less likely to see a signed review for major revisions than for minor revisions, providing further support for the generalizability of the observed effect.*

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.

**We have access to reviewer names only when they signed. Therefore, we do not know how often the same person signed their review, and how often they did not. For example, in our dataset, we see that Daniel Lakens signed reviews 3 times for PeerJ, and once for Royal Society**

Open Science. But our data does not allow us to draw conclusions about how often Daniel Lakens did not sign, and if he selectively signed from some papers but not for others. We see most reviewers appear once, 275 appear twice, and a smaller set have 3 or more reviews. But again, there is no way to know if they always sign, or not. We have noted this in the discussion.

review\_names\_peer\_j

1 time: 3815

2 times: 275

3 times: 54

4 times: 9

5 times: 2

6 times: 2

7 times: 1

review\_names\_royal\_society

1 time: 1410

2 times: 54

3 times: 5

4 times: 4

We see most reviewers appear once in the first round of review of a manuscript, some appear twice, and a smaller set have 3 or more reviews. But again, there is no way to know if they always sign, or not. We have noted this in the discussion.

*It might be that some reviewers consistently sign their reviews, while others sign selectively. Some researchers have contributed multiple signed reviews to our dataset, but it is unknown whether they also contributed unsigned reviews. The confidentiality of peer review makes it impossible to retrieve whether reviewers selectively sign from open reviews, and this question will need to be studied by surveying researchers directly.*

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.

**Spoiler:** People have not been persuaded. To address point 4 of reviewer 1, about changes over time, we updated our textmining script for PeerJ to also retrieve the date the paper was accepted for publication, as well as the manuscript script where we retrieved publication year from the scopus file for Royal Society Open journals. We have added a figure to the main article that shows the recommendations as a function of signed or non-signed reviews per year, because it seems at least some readers (such as Reviewer 1) would believe signed reviews are becoming more common. However, as the data shows, this is not the case. We already wrote in the previous version (and have not changed the sentence) that: “The percentages of people who signed (38.23% for PeerJ, 36.99% for Royal Society Open journals) are slightly lower than the 43.23% reported by Wang, You, Manasa, and Wolfram (2016) who analyzed the first 1214 articles published in PeerJ”. In other words, there is actually a decline in the percentage of signed reviews, and the new figure makes this even more clear. One can only speculate why this is the case, but one reason might be that early adopters (those academics who reviewed for RSOS and PeerJ when the journals were new) were more likely to sign, but as the journal became wider known, and received more submissions, the journals relied on more the more average scientist, who might be less likely to sign than early adopters. Whatever the reason, it is informative to know progress is not as imminent as Reviewer 1 expects.

#### 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.

**Reviewer 1 notes that perhaps RSOS and PeerJ ask authors to recommend reviewers. We checked, and this is indeed the case.**

**RSOS has in the submission portal the following instruction:**

*You are required to supply at least 4 preferred referees, and are required to supply a reason for the preference. Enter the reviewer's information into the textboxes below and click the appropriate designation button.*

*Royal Society Open Science is committed to increasing the diversity of its reviewer pool. When recommending reviewers, we would particularly appreciate suggestions of suitably qualified researchers from underrepresented groups (including women, ethnic minority scientists, scientists with disabilities and other underrepresented groups), early career researchers, and researchers from the global South.*

*Your suggestions should:*

- include an institutional email address (ie not a yahoo, gmail or equivalent commercial domain) to help us verify their identity;*
- be experts in the field (ie at least a PhD and an established, verifiable publication record);*
- not be someone with whom you have a financial, professional, or personal conflict of interest. Conflicts of interest include, but are not limited to working or having recently worked in the same institution or department as any of the authors and having recently (eg in the past 3 years) been a supervisor, mentor, mentee, close collaborator or joint grant holder with any of the authors.*

*If you are indicating opposition to a possible referee, please provide a full reason for the opposition – an explanation of the nature of the conflict is required if you state you have a conflict of interest with a specific referee. It is insufficient to simply state you have a conflict.*

*To designate preferred and non-preferred editors, select them from the dropdown and click the appropriate designation button.*

PeerJ has the following instructions in the submission portal:

*2. Suggest at least 2 editors, and 3 reviewers for your article. Conflict of Interest checks:*

*Reviewers must not have published with you or co-authors in last 5 years. Reviewers must not be at the same institution as you or co-authors. Your submission may be rejected if a Conflict of Interest is identified.*

**There is no way to know which reviewers in our dataset were suggested and which were not, nor what the impact of this is on how often they sign. If anything, personally we would believe a main effect might be most plausible (more positive recommendations by suggested reviewers) but it is possible that suggested reviewers are more likely to not sign when giving negative recommendations than non-suggested reviewers. Therefore, we added a sentence indicating that 1) Royal Society Open journals and PeerJ allow authors to recommend reviewers, and 2) that we do not know if our findings will generalize to journals that have different policies:**

*It is possible to indicate suggested and opposed reviewers when submitting a manuscript at PeerJ and RSOS. It is unknown how often suggested reviewers end up reviewing submissions, or whether suggested or non-suggested reviewers differ in how often they sign their reviews, depending on their recommendation. It therefore is unknown whether the current results will generalize to journals that do not ask authors to recommend reviewers.*

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.

Regarding comment 6, the reviewer raised a very good point about there being multiple reasons not to sign, and how we highlighted only one. Although we think negative backlash is a clear concern based on additional research, we have added a sentence in the discussion highlighting different reasons for not signing, which included the concern about CoI's raised by the reviewer:

*Reviewers might have different reasons to sign or not sign their reviews. They might be worried about writing a positive review for a paper that is later severely criticized, and receiving blame for missing these flaws when reviewing the submission. Reviewers might not want to reveal how their recommendations are unrelated to the scientific quality of the paper, but based on personal biases or conflicts of interest. Finally, reviewers might fear backlash from authors who receive their negative reviews.*

We have corrected all spelling errors (also those kindly pointed out by the editor) and the link to the Bastian blog post should work now.

## Reviewer 2

1: 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.

**We have added some information about the scope of PeerJ and Royal Society Open journals, in the section in the current study, which already contained important information about the journals:**

*PeerJ and Royal Society Open journals publish articles across a wide range of scientific disciplines, including biology, chemistry, engineering, life sciences, mathematics, and medicine*

**Concerning the reputation of these journals:** We do not feel qualified to make a judgment on behalf of others. We think these are good journals, and have published in both, and reviewed for both. But researchers should make up their own mind, and the experiences researchers have might differ across scientific disciplines, so we stick with objective facts when describing the journal (e.g., how long it has had open reviews, how long the journals exist, and the topics they publish on). We also discuss the scope of journals in the new Figure plotting the data across sections in PeerJ.

2: 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.

**We have added titles to the figures to make the figures more distinct.**

3: \*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.

**We understand this comment by the reviewer, and agree the 'additional analyses' header immediately after the 'results' section was an unfortunate choice. What we wanted to achieve was to highlight the richness of the datafile we are sharing, and invite fellow meta-scientists to make use of this data. We also made this point in the general discussion, and have now removed the 'additional analyses' section, and instead discussed the variables in the dataset, including the analysis of agreement among reviewers as an example, in the discussion. The section now follows the sentence in the first version of the manuscript: "Open reviews allow researchers to examine meta-scientific questions that give insights into the peer review process." and is followed by "We hope the dataset we share with this manuscript will allow other meta-scientists to examine additional questions about the peer review process." We hope this makes the point of this section clearer (it is not presented as an analysis related to the main question in our**

manuscript, but an illustration of the kind of questions that could be answered in our data analysis).

Point 4: The reviewer raises a very important point how reputations could easily be damaged by this type of transparency if it became known one reviewed retracted papers. We have combined this with addressing point 6 by reviewer 1 and added the following to the discussion:

*Reviewers might have different reasons to sign or not sign their reviews. They might be worried about writing a positive review for a paper that is later severely criticized, and receiving blame for missing these flaws when reviewing the submission. Reviewers might not want to reveal how their recommendations are unrelated to the scientific quality of the paper, but based on personal biases or conflicts of interest. Finally, reviewers might fear backlash from authors who receive their negative reviews.*

5: 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.

**We already raised a counter narrative in the results section, but perhaps the reviewer means we should also highlight the opposite relationship in the discussion, which is a fair comment that we incorporated in the revision. We have now added in the discussion:**

*Due to the correlational nature of the data, this relationship could emerge because reviewers who are on average more likely to sign also give more positive recommendations, reviewers are more likely to sign more positive recommendations, or both these effects could be true at the same time.*

## Editor comments:

2: 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.

We do not present quantitative analyses because we have analyzed the entire population of open reviews in PeerJ and Royal Society Open journals, and are simply presenting and interpreting the descriptive statistics. Since our main results are presented as proportions, we doubt recalculating these as odds ratios adds much about the magnitude of the effect - we hope proportions are relatively straightforward to understand. We could add an odds ratio for the results for PeerJ (but remember we have excluded 22 directly accepted results, so there is a third category), but this would be more complex for Royal Society Open journals given the 4 recommendation categories. We also wouldn't know which effect sizes in the literature to compare the effect sizes with, and given main effects between journals (e.g., differences in the proportion of minor vs major revision recommendations) we would not know what analyses to report, beyond providing researchers with the descriptives based on the entire population. We are happy to consider any specific suggestions.

3: 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.

This is indeed an interesting point to mention. Reviewers do not see the recommendation (although reviewers make a recommendation). We have included this information in our discussion about how PeerJ and Royal Society Open journals work. We hesitate to discuss possible consequences of this. One might believe it would make reviewers more comfortable to recommend to reject, even while signing, because they can 'hide' behind the editor decision. Reviewers still get to read the reviews. It could be that the private decision differs from the publicly available review. But there is no way to know without having access to confidential information about the peer review process.

4: "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.

*We agree - This comment aligns with that of Reviewer 1, point 1 and 2, and our new table for PeerJ which shows the same pattern across all disciplines suggest the pattern we have observed is not specific to any field, but surprisingly general.*

5: “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.

**This is a good point that we missed, but which is echoed in comment point 4 by reviewer 2, and point 6 by Reviewer 1. We have addressed this in the new sentences in the discussion:**

*They might be worried about writing a positive review for a paper that is later severely criticized, and receiving blame for missing these flaws when reviewing the submission. Reviewers might not want to reveal how their recommendations are unrelated to the scientific quality of the paper, but based on personal biases or conflicts of interest. Finally, reviewers might fear backlash from authors who receive their negative reviews.*

6: 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.

**You probably ran the code, line by line. If you ‘knit’ the document, the papaja package will reproduce the submitted PDF exactly, including the rounding. We will make sure that our manuscript is available as a completely reproducible Code Ocean capsule, so that anyone can reproduce our results, without needing to install any software locally.**