Researchers Decision to Sign Reviews is Related to Their Recommendation

Nino van Sambeek[1] & Daniël Lakens[1]

[1] Eindhoven University of Technology, The Netherlands

Surveys indicate that researchers are generally positive about making peer reviews available alongside published articles. Researchers are more negative about revealing the identity of reviewers, They worry reviewers will be less likely to express criticisms if their identity is known to authors. Experiments suggest that reviewers are somewhat more likely to recommend to accept articles when they are told their name will be shared alongside the reviews, than when they will remain anonymous. One recent study revealed reviewers in five journals gave more positive recommendations when during a peer review experiment their identity would be made available to authors. We replicate and extend this finding by analyzing 12010 open reviews in PeerJ and 4188 reviews in the Royal Society Open Science and Open Biology where authors can voluntarily make their identity known. Our results confirm that the probability that a reviewer reveals their identity is related to their recommendation. The more positive a recommendation was, the more likely it was that authors shared their identity. Although our data is correlational and does not allow causal conclusions, it provides behavioral data from real peer reviews that demonstrates convincingly that reviewers who sign their reviews give more positive recommendations. We also share all 23649 text mined reviews as raw data underlying our results that can be re-used by researchers interested in peer review.

Keywords: Peer Review, Open Science, Transparency
Word count: 2711

As technology advances, science advances. The rise of the internet has made it possible to share all aspects of the scientific process. This includes opening up the peer review process. An increasing number of journals has started to make peer review reports available alongside published articles as part of ongoing experiments that aim to improve the peer review process (Bruce, Chauvin, Trinquart, Ravaud, & Boutron, 2016). Open peer review can be implemented by making peer reviews available, but also be revealing the identity of reviewers during or after the peer review process. An important argument for revealing the identity of reviewers is that they can receive credit for their work (Godlee, 2002). However, scientists do not feel these benefits outweigh possible costs, and are worried that criticism on manuscripts might lead to backlash from the authors in the future. Some reviewers might accept these consequences, while other might choose to strategically reveal their identity only for positive reviews they write.

Researchers indicate that they would be less likely to review for a journal if their names are made public, and anecdotally mention that signed reviews would make it more difficult to be honest about manuscripts they believe are poor quality (Mulligan, Hall, & Raphael, 2013). A more recent survey found that 50.8% of almost 3000 scientists believe that revealing the identity of reviewers would make peer review worse (Ross-Hellauer, Deppe, & Schmidt, 2017). Almost two-thirds of respondents believed reviewers would be less likely to deliver strong criticisms if their identity became known to the authors.

These self-report studies are complemented by experiments in which reviewers are randomly assigned to a condition where their identity would be revealed during the peer review process (Walsh, Rooney, Appleby, & Wilkinson, 2000). Reviewers in the condition where their identity was revealed were less likely to recommend rejection (n = 30) than reviewers who remained anonymous (n = 51). This suggests that a causal effect exists between knowing your identity will be revealed, and the recommendation that is made during the peer review process. A meta-analysis by Bruce and colleagues (2016) supports the conclusion that across four studies reviewers are somewhat less likely to recommend rejection when they have to sign their reviews.

Although the self-report studies and the experiments clearly suggest reviewers worry about having their name attached to more critical reviews they write, so far little is known about what reviewers actually do when given the opportunity to sign their reviews. The trade-off between the benefit of

This work was supported by the Netherlands Organization for Scientific Research (NWO) VIDI grant 452-17-013. All data underlying this reproducible manuscript are available at https://osf.io/jkbmw/

Correspondence concerning this article should be addressed to Daniël Lakens, ATLAS 9.402, 5600 MB, Eindhoven, The Netherlands. E-mail: D.Lakens@tue.nl

getting credit for peer review work and the risk of negative consequences when signing critical reviews might lead to strategic behavior where authors become more likely to sign reviews the most positive their recommendation is. If this strategic behavior occurs in practice, we should see a different pattern of recommendations for signed and unsigned reviews. One recent study revealed such a pattern when analyzing data from an Elsevier trial on publishing peer review reports in the journal Agricultural and Forest Meteorology, Annals of Medicine and Surgery, Engineering Fracture Mechanics, Journal of Hydrology: Regional Studies, and the International Journal of Surgery (Bravo, Grimaldo, López-Iñesta, Mehmani, & Squazzoni, 2019). Although only 8.1% of reviewers choose to disclose their identity in these reviews, the data revealed a clear difference between the recommendations by reviewers who chose to sign their reviews, compared to reviewers who did not choose to sign.

The Current Study

We examined the relationship between recommendations peer reviewer make and the probability that they have signed their review in two large open access journals, PeerJ and Royal Society Open Science and Open Biology. These journals publish across a wide range of scientific disciplines, thus allowing us to replicate and extend the analysis by Bravo and colleagues (2019). To examine the pattern of recommendations as a function of whether reviewers signed their review or not we analyzed 7930 articles published in PeerJ, as well as 3576 articles published in The Royal Society Open Science (RSOS) and Royal Society Open Biology (RSOB). PeerJ launched in 2012 and provides reviewers the possibility to sign, and provides authors the possibility to make peer reviews available with the final publication. RSOS launched in 2014 and strongly encouraged authors to make the peer reviews available with the final publication, and made this mandatory in January 2019. RSOB made sharing reviews with the final publication mandatory in May 2017. Peer reviewers have the option to make their identity known when submitting their review to RSOS or RSOB. Because of their broad scope, the large number of publications in each journal, and their early focus on open reviews, these reviews provide insights into the peer review behavior of scientists across a wide range of disciplines.

Accessing Open Reviews

PeerJ assigns all articles a number, increasing consecutively with each published manuscript. Reviews are always accessible in html (i.e., at https://peerj.com/articles/1/reviews for the first article can be found). PeerJ has several more recent dedicated journals for computer science and chemistry. In the analyses reported here we included PeerJ Computer Science, which started in 2015, but ignore the chemistry journals that started at the end of 2019 due to the small number of published articles. For RSOS and RSOB reviews are published online as a PDF file. A list of Digital Object Identifiers (DOIs) for every article published in these two outlets was retrieved through Scopus. All reviews were downloaded, and the PDF files were converted to plain text files using pdftools for R (Ooms, 2019; R Core Team, 2013). These text files were mined for recommendations, reviewers names, submission and acceptance dates, and the review content, using the stringr package in R (Wickham, 2019).

For each article we extracted the number of revisions, and for each revision we saved whether each of the reviewers signed, the word count for their review, and their recommendation for that review round. Note that for PeerJ the editor makes the recommendation for each revision, based on the reviews. We therefore do not directly know which recommendation each reviewer provided, but analyze the data based on the assumption that the decision by the editor is correlated with the underlying reviews. For RSOS and RSOB reviewers do recommend to "accept as is", "accept with minor revisions", "major revision", or "reject". Although "reject" recommendations occur for RSOS and RSOB, both PeerJ, RSOS, and RSOB only share reviews for published articles, and therefore we have very few "reject" recommendations in RSOS and RSOB, and none in PeerJ. Searching all reviews for PeerJ for the words "appealed on" revealed 44 articles that were initially rejected, appealed, received a "major revision" recommendation, and were eventually published. We have coded these papers as "major revisions". All scripts to download and analyze the reviews, and computationally reproduce this manuscript, are available at https://osf.io/jkbmw/.

Results

We examined 8155 articles published in PeerJ (7930 in PeerJ, 225 in PeerJ Computer Science), as well as 3576 articles from Royal Society (2887 from RSOS, 689 from RSOB, 81 of which were editorials or errata) published up to October 2019. We retrieved all reviews when these were made available (5087 articles in PeerJ, 1964 articles in Royal Society). Articles can go through multiple rounds of review. We focus on the first review round as this review reflects the initial evaluation of reviewers, before the handling editor has made a decision, following Bravo et al. (2019). On average initial submissions at PeerJ received 2.36 reviews. Articles in the Royal Society journals received on average 2.13 reviews for the original submission.

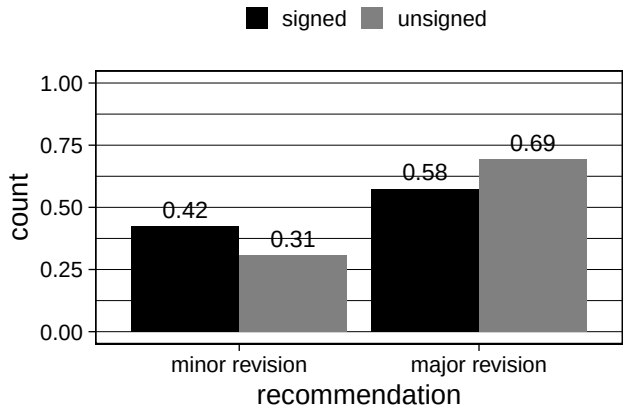

Figure 1. Proportions of signed and unsigned reviews as a function of whether the handling editor at PeerJ recommended "minor revisions" or "major revisions".

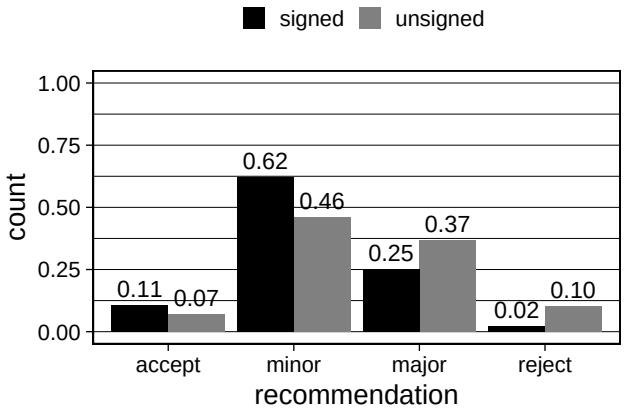

Figure 2. Proportions of signed and unsigned reviews as a function of whether reviewers at Royal Society Open Science or Open Biology recommended "accept", "minor revisions", "major revisions", or "reject".

Signed reviews as a function of the recommendation

For all 5087 articles published in PeerJ where reviews were available we retrieved 12010 unique reviews (as each article is typically reviewed by multiple reviewers) for the initial submission. In total 4592 reviewers signed their review for the initial submission, and 7418 reviewers did not. In Royal Society journals we analyzed 1964 articles for which we retrieved 4188 unique reviews for the first submission, where 1549 reviewers signed their review, and 2639 did not. The percentages of people who sign (38.23% for PeerJ, 36.99% for Royal Society) are slightly lower than the 43.23% reported by (Wang, You, Manasa, & Wolfram, 2016), who analyzed the first 1214 articles published in PeerJ.

To answer our main research question we plotted the signed an unsigned reviews as a function of the recommendation in the first review round (see Figure 1). Remember that for PeerJ these recommendations are made by the editor, and thus only indirectly capture the evaluation of the reviewer. For minor revisions, a greater proportion or reviews was signed than unsigned, but for major revisions, more reviews were unsigned than signed. Too few articles are immediately accepted after the first round of reviews in PeerJ (22 in total) to impact the proportions in the other two categories.

Analyzing the reviews at the Royal Society provides a more direct answer to our question, since each individual reviewer is asked to provide a recommendation of "accept", "minor revisions", "major revisions", or "reject". We can therefore directly compare how recommendations are related to the decision to sign reviews (see Figure 2). The overal pattern clearly shows that the proportion of signed reviews is larger for more positive recommendations (accept and minor revisions) whereas the proportion of unsigned reviews is larger for more negative reviews (major revisions and reject).

We can not draw causal conclusions based on this correlational data. It is possible that reviewers are less likely to sign more negative reviews. It is also possible that people who sign their reviews generally give more positive recommendations, and therefore the distribution of signed reviews differs from non-signed reviews. Based on the literature reviewed in the introduction we know researchers are hesitant to voice criticism when their identity will be known, and experimental evidence suggests that if identities are shared with authors, recommendations become somewhat more positive. Therefore, it seems plausible that at least part of the pattern we observed can be explained by reviewers being more likely to sign their more positive reviews. Although we had access to few "reject" recommendations because we only analyzed reviews for published manuscripts, the difference between signed and unsigned reviews for major revisions, minor revisions, and accept recommendations replicates the findings by Bravo et al. (2019) across a larger range of research fields, based on a larger dataset, and in journals where a larger percentage of reviewers volunteers to disclose their identity. This replication suggests the difference in recommendations depending on whether reviews are signed or not is a reliable observation.

Additional Analyses

The dataset we are sharing has information about the recommendations of reviewers (RSOS and RSOB) or editors (PeerJ) after each round of peer review, the names of reviewers who signed their review, and the time in review (114 days for PeerJ, 132 days for RSOS and RSOB). Through the DOI, researchers can link this data to other sources of information such as ci-

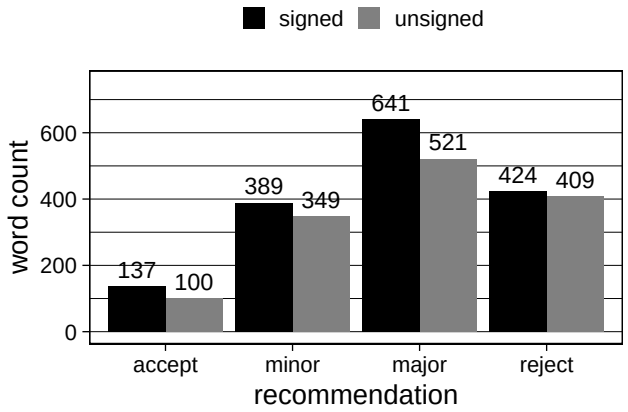

Figure 3. Word count for reviews for first submission at Royal Society journals as a factor of the recommendation.

tation counts. Because the reviews themselves are included in our dataset, researchers can use the text files to answer more detailed questions about the content of peer reviews across different domains. For example, we can examine the word count for signed and unsigned reviews as a function of the recommendation. With the caveat that the graph below present uncleaned raw word counts, we replicate the results reported by (Wang et al. (2016)) for PeerJ in the dataset for Royal Society Open Science and Open Biology that there is no statistical difference between the word count for signed and unsigned reviews, despite a small numerical difference in the same direction as in their analysis of PeerJ reviews (see Figure 3).

Since we know the individidual recommendation of each reviewer for Royal Society journals, one example of the insights open reviews provide is how often reviewers agree. For the 1963 papers where te reviews were published, all reviewers agreed on the recommendation for 830 articles (s0 42.28% of the time). For 41.72% of the manuscripts the maximum deviation was one category (e.g., minor and major revisions), for 14.37% of the manuscripts the maximum deviation was two categories (e.g., accept and major revision), and for 1.63% of the manuscripts the maximum deviation was three categories (i.e., accept and reject). There were 3 articles where researchers received all four possible recommendations (accept, minor revisions, major revisions, reject) from at least four different reviewers.

Regrettably, neither PeerJ nor RSOS make peer reviews available for manuscripts that were rejected. As a consequence, we have analyzed a biased sample of the literature. At the moment only very few scientific journals (e.g., Meta-Psychology, F1000) make peer reviews available for all submitted articles. Although open reviews enable us to look in more detail at the peer review process, it would be extremely interesting to be able to follow manuscripts through the peer review process even when they are rejected at one specific journal. Despite this limitation, the pattern of results we observe is very similar to that reported by Bravo et al. (2019) who had access to the reviews for accepted and rejected manuscripts. Thus, even though the proportion of reject recommendation in our dataset is small, the relative difference with which this recommendation is given in signed and unsigned reviews seems comparable to datsets that include rejected manuscripts.

Discussion

Our analysis shows that when authors are given the choice to sign their reviews, signed reviews have more positive recommendations than unsigned reviews. This pattern is clearly present for reviews in Royal Society Open Science, a large multi-disciplinary journal that published across a wide range of scientific domains. The pattern is also visible in a second large multi-disciplinary journal, PeerJ, under the assumption that recommendations by editors at PeerJ are correlated with the recommendations by reviewers. Our results replicate and extend earlier findings by Bravo et al. (2019), and complement self-report and experimental results in the literature.

Peer review is generally seen as an important quality control mechanism in science, yet researchers can rarely evaluate the quality of peer review at journals. Open reviews allow researchers to examine meta-scientific questions that give insights into the peer review process. Our data support the idea that researchers decision to sign is related to their recommendation across a wide range of scientific disciplines. Together with self-report data and experiments reported in the literature, our data increase the plausibility that in real peer reviews at least some researchers are more likely to sign if their recommendation is more positive. This type of strategic behavior also follows from a purely rational goal to optimize the benefits of peer review while minimizing the costs. 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.

It is worthwhile to examine whether this fear of retaliation has an empirical basis, and if so, to consider developing guidelines to counteract such retaliation (Bastian, 2018). Although the available data suggests that at least some reviewers hesitate to sign if they believe doing so could have negative consequences, but will sign more positive recommendations to get credit for their work, it seems worthwhile to explore ways in which reviewers feel comfortable to claim credit for all their, regardless of whether their recommendation is positive or negative.

Author Contributions

N. van Sambeek and D. Lakens developed the idea, and jointly created the R code to generate and analyze the data. N van Sambeek drafted the initial version of the manuscript as a Bachelor thesis, D. Lakens drafted the final version, and both authors revised the final version of the manuscript. The authors report no conflicts of interest.

References

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/.

Bravo, G., Grimaldo, F., López-Iñesta, E., Mehmani, B., & Squazzoni, F. (2019). The effect of publishing peer review reports on referee behavior in five scholarly journals. *Nature Communications*, *10*(1), 1–8. doi:10.1038/s41467-018-08250-2

Bruce, R., Chauvin, A., Trinquart, L., Ravaud, P., & Boutron, I. (2016). Impact of interventions to improve the quality of peer review of biomedical journals: A systematic review and meta-analysis. *BMC Medicine*, *14*. doi:10.1186/s12916-016-0631-5

Godlee, F. (2002). Making Reviewers Visible: Openness, Accountability, and Credit. *JAMA*, *287*(21), 2762–2765. doi:10.1001/jama.287.21.2762

Mulligan, A., Hall, L., & Raphael, E. (2013). Peer review in a changing world: An international study measuring the attitudes of researchers. *Journal of the American Society for Information Science and Technology*, *64*(1), 132–161. doi:10.1002/asi.22798

Ooms, J. (2019). *Pdftools: Text extraction, rendering and converting of PDF documents*.

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ross-Hellauer, T., Deppe, A., & Schmidt, B. (2017). Survey on open peer review: Attitudes and experience amongst editors, authors and reviewers. *PLOS ONE*, *12*(12), e0189311. doi:10.1371/journal.pone.0189311

Walsh, E., Rooney, M., Appleby, L., & Wilkinson, G. (2000). Open peer review: A randomised controlled trial. *The British Journal of Psychiatry*, *176*(1), 47–51. doi:10.1192/bjp.176.1.47

Wang, P., You, S., Manasa, R., & Wolfram, D. (2016). Open peer review in scientific publishing: A Web mining study of PeerJ authors and reviewers. *Journal of Data and Information Science*, *1*(4), 60–80.

Wickham, H. (2019). *Stringr: Simple, consistent wrappers for common string operations*.