Are peer reviewers influenced by their work being cited?

Adrian Barnett [1,*]

[1]School of Public Health & Social Work, Queensland University of Technology, 88 Musk Avenue, 4059, QLD, Australia
[*]a.barnett@qut.edu.au

October 1, 2025

Abstract

Peer reviewers sometimes comment that their own journal articles should be cited by the journal article under review. Comments concerning relevant articles can be justified, but comments can also be unrelated coercive citations. Here, we used a matched study design to explore how citations influence the peer review process. We used a sample of more than 37,000 peer reviews from four journals that use open peer review and make all article versions available. We find that reviewers who were cited in the article under review were more likely to recommend approval, but only after the first version (odds ratio = 1.61; adjusted 99.4% CI: 1.16 to 2.23). For all versions of the articles, the reviewers who commented that their own articles should be cited were less likely to recommend approval compared to the reviewers who did not, with the strongest association after the first version (odds ratio = 0.15; adjusted 99.4% CI: 0.08 to 0.30). Reviewers who included a citation to their own articles were much more likely to approve a revised article that cited their articles compared to a revised article that did not (odds ratio = 3.5; 95% CI: 2.0 to 6.1). Some reviewers' recommendations depend on whether they are cited or want to be cited. Reviewer citation requests can turn peer review into a transaction rather than an objective critique of the article.

Introduction

In 2024, a published peer reviewed article included this remarkable sentence: "As strongly requested by the reviewers, here we cite some references [35–47] although they are completely irrelevant to the present work" [1]. This was a rare public example of coerced citations, where a reviewer exploits the peer review process to increase their citation counts and hence further their own career [2–4]. Reviewers should be relevant experts, so some suggestions to cite their articles will be appropriate. However, excessive citation requests or requests to cite unrelated articles are unethical [5–9]. Coerced citations can also come from editors trying to boost their journal's ranking [10–12].

Coerced citations are reported as a common problem in peer review. In author surveys, two-thirds reported pressure from peer reviewers to cite unrelated articles [13] and 23% had experienced a reviewer that "required them to include unnecessary references to their publication(s)" [14]. Publishers have investigated whether "hundreds of researchers" have manipulated the peer review process to increase their own citations [15]. Some reviewers may be exploiting their power over authors who "have a strong incentive to [...] accept all 'suggestions' by the referees even if one knows that they are misleading or even incorrect" [16].

As reviewers are often in the same field as the article's authors, they may already be cited in the article without the need for coerced citations. Reviewers who are cited may give a more favourable peer review and be more willing to overlook flaws [17, 18]. Some authors may try to exploit this using "referee baiting" [3] or "flattery citations" [19] by favourably citing a reviewer's work.

The interactions during peer review between authors and reviewers can determine whether an

article is accepted [20] and what results are included in the published version [21]. Given the importance of peer review for science, studies that examine how peer review works in practice are needed [22–26]. Here, we examine interactions between peer reviewers and authors using four journals that publish all article versions and all peer reviews. We had two research questions:

1. Do peer reviewers give a more or less favourable recommendation when they are cited in the article?
2. Do peer reviewers give a more or less favourable recommendation when their review includes a citation to their own articles?

Methods

Journal selection

We studied journals from the publisher *F1000* as their journals use open peer review with signed reviewers. *F1000* journals use a publish–review–curate model [27], meaning all versions of the article are publicly available, including versions updated after peer review. This allowed us to examine the interactions between authors and reviewers throughout the peer review process. We selected four *F1000* journals that each had over 100 articles. Some characteristics of the four journals are given in Table 1. Three journals were created to support funders.

Table 1: **Brief information about the four included journals from the publisher *F1000*.**

Journal title	Year started	Field(s) of research	Articles must concern research funded by
F1000Research	2012	All disciplines	*No restriction*
Wellcome Open Research	2016	Medicine, Genomics	Wellcome
Gates Open Research	2017	Medicine	The Gates Foundation
Open Research Europe	2021	All disciplines	European Commission

The peer review process used by *F1000* journals differs from most standard journals. The journals do not use academic editors, but do have in-house editors who manage articles but do not make editorial decisions. This means that most interactions during peer review are between authors and reviewers directly. In-house editors perform checks prior to the first version of the article being published and at *F1000Research* this results in 40 to 50% submissions being rejected (personal communication, *F1000* staff). Up to mid-2024, authors were asked to identify potential reviewers who were qualified experts with no competing interests [28]. Since mid-2024, reviewer identification is made in-house, although authors can suggest reviewers.

Reviewers are asked to recommend one of three categories: Approved, Approved with reservations, and Not approved. For brevity, we refer to 'Approved with Reservations' as 'Reservations'. An article is indexed once it receives two 'Approved' or two 'Reservations' and one 'Approved'. The guidelines for recommending Approved are: "the aims and research methods are adequate; results are presented accurately, and the conclusions are justified and supported by the presented data" [29]. Peer reviewers are asked to assess the validity of an article's content, rather than novelty or interest levels, an approach designed to combat publication bias [30].

All four journals have a peer reviewer code of conduct and state that reviewers should familiarise themselves with the ethical guidelines for peer reviewers by the *Committee On Publication Ethics* [31]. The journals' guidelines for reviewers include the following: "reviewers should explicitly state their reasoning when asking authors to cite their own work".

Data extraction

We extracted data on authors and articles from the *OpenAlex* database (`https://openalex.org/`) and directly from the four journals. *OpenAlex* combines scholarly data from multiple sources, including *ORCID* – a unique identifier for researchers, *Microsoft Academic*, *Crossref* and *PubMed*. A recent study compared *OpenAlex* with the two commonly used proprietary bibliometric databases of *Web of Science* and *Scopus* for the years 2015 to 2022 [32]. The results were mixed, but *OpenAlex* had better *ORCID* coverage and covered more

Digital Object Identifiers (DOIs) – the unique identifier for publications. We accessed *OpenAlex* using the *openalexR* package [33, 34]. We used each journal's application programming interface (API) to extract data on the articles and peer reviews. The data were extracted in four stages:

1. Searches were made using the APIs of the four journals to find all articles published between 1 Jan 2012 and 28 May 2025, with the start date to capture all potential articles.
2. For each article, the following data were downloaded in XML format:
    - The article's publication date and version number
    - The reviewers' names and *ORCIDs* (if available)
    - The text of all reviews and the reviewers' recommendations
    - The DOIs and PMIDs (*PubMed* IDs) from the article's reference list
    - The DOIs and PMIDs of any articles cited by the reviewers. The online peer review system at *F1000* journals includes the DOI of any article cited in the review, which facilitates the identification of citations to the reviewers' articles.
3. Articles were excluded if:
    - They were not peer reviewed or had yet to receive any reviews
    - The reference list was empty
4. The reviewers' publication histories were collected from *OpenAlex* using their name, institution and *ORCID* (if available). Reviews were excluded if there was no record for the reviewer in *OpenAlex*, or if the reviewer had no published articles as there was no potential for them to be cited or request a citation to their own articles.

Study design

We used two predictor variables about the reviewer:

- The number of times they were cited in the article $(0, 1, 2, \ldots)$.
- The number of times they included citations to their own articles in their review $(0, 1, 2, \ldots)$.

We fitted both predictors as linear, but reviewers may behave differently with any citation rather than a linear change, and hence we also fitted both predictors as a binary "none versus any" (0 versus $1, 2, \ldots$). We compared the linear and binary alternatives using the Akaike Information Criterion (AIC) to find the parameterisation that best fitted the data [35].

We matched on article and version to control for confounding by any characteristics of the article [36]; for example, the article's topic or writing style. Hence, we compared two or more independent reviewers who considered the same article.

All analyses were stratified by article version, using the first version only or the second and subsequent versions. This is because the reviewers are unknown to the authors for the first version, but from the second version onwards, the authors will know the reviewers as the journals use signed peer reviews. This knowledge could alter the behaviour of authors and reviewers.

The study design is summarised in Figure 1

Statistical methods

We used conditional logistic regression to examine the associations between the citations to the reviewer and their ordinal recommendation (Approved → Reservations → Not approved) while matching the article and the version [37]. Conditional logistic regression requires a binary dependent variable; hence, we fitted two related models that examined the odds of:

1. "Approved" compared with "Reservations" or "Not approved".
2. "Approved" or "Reservations" compared with "Not approved",

These two models tested the same hypothesis, hence we adjusted for multiple testing. We also used repeated testing due to the stratification by article version and the two formulations of the predictors (linear or none versus any). Since we used 8 ($2 \times 2 \times 2$) tests, we displayed all the results using 99.4% confidence intervals instead of 95.0% intervals, which is a 5% type I error divided by eight tests.

In an unplanned analysis, we examined the association between the reviewer's recommendation and whether they included citations to work other than citations to their own articles. This was added to examine differences between reviewers' citations to their own articles and other articles.

Outliers were not excluded. No data were missing in the analysis data set.

The sample size calculation is in Supplement S.1.

Text analysis

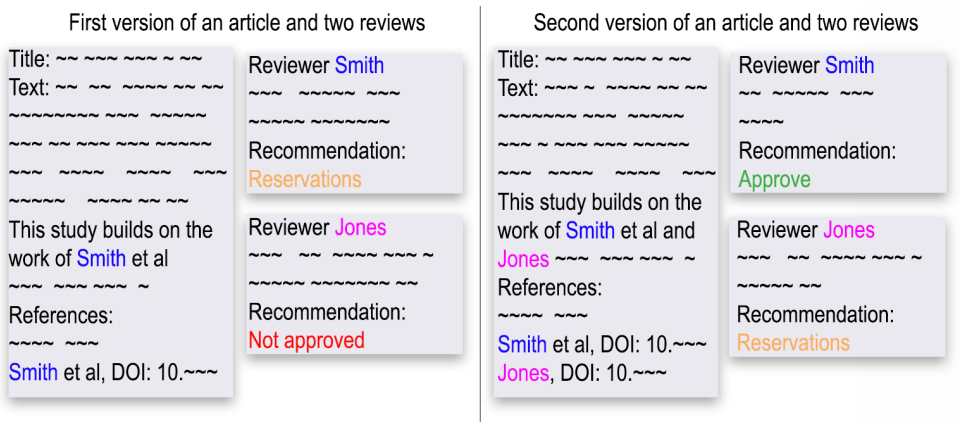
We examined how reviewers' citations to their own articles or other articles were justified and whether their wording differed according to their recommendation. For an initial view of citations to their own articles, we randomly selected 20 reviews and extracted the most relevant sentence concerning the citation.

To analyse the review text, we first extracted the 100 most commonly used words in all reviews. To standardise the text, all words were transformed into tokens, with stop-words removed and then stemmed. We then tested which of the 100 words were associated with recommending Approved versus Reservations or Not approved amongst those reviewers who included a citation to their own articles and those who included a citation to other articles. We chose the set of words using an elastic net with 10-fold cross-validation and selected a parsimonious model by using the lambda within one standard error of the minimum cross-validated error [38, 39]. To get uncertainty intervals for the estimates, we fitted a Bayesian model with the set of words selected by the elastic net and using a sceptical Normal prior centred on zero to create shrinkage.

Reproducibility

Research question 1 was pre-registered using *As Predicted* on 20 May 2024 [40]. Research question 2 was formulated during data collection but before any data analysis and used the same study design and statistical methods as question 1.

Figure 1: **Graphical summary of the study design for research question 1 showing a dummy article and two reviews.** *In the first version of the article, the reviewer Smith (blue) is cited whilst Jones (purple) is not. For the second version of the article, the authors are now aware that Jones is a reviewer and Jones has been cited. The reviewers' recommendations are the outcome and are colour-coded as Not approved (red), Reservations (orange) and Approved (green). We tested whether citations to the reviewer in the article influenced their recommendation. The matched design means that only reviewers of the same article are compared (here, Smith and Jones) and the overall effect is estimated by aggregating over multiple matched comparisons. Research question 2 used the same design but examined citations to the reviewers' articles in their reviews.*



All data extraction and analyses were conducted using *R* version 4.4.1 [41]. The data and *R* code are available on *GitHub* [42].

Results

A flow chart of the included reviews is shown in Supplement S.2. The final sample size was over 37,000 reviews. There were more than 3,500 articles that were not included because they had not yet been peer reviewed, especially recent articles. More than 2,000 reviewers did not have a record in *OpenAlex* and so could not be included. These missing reviewers were more likely to be from older articles and more likely to be co-reviewers.

Descriptive statistics on the included reviews are in Table 2. The reviewers were cited at least once in 13% of the articles and 6% of the reviews included a self-citation. Most reviews recommended "Approved" (54%), with only 8% recommending "Not approved" which is low compared with many journals; however, 40 to 50% of submissions are rejected before articles are sent for peer review (personal communication, *F1000* staff). The reviewers were relatively experienced, with a median number of papers of 55.

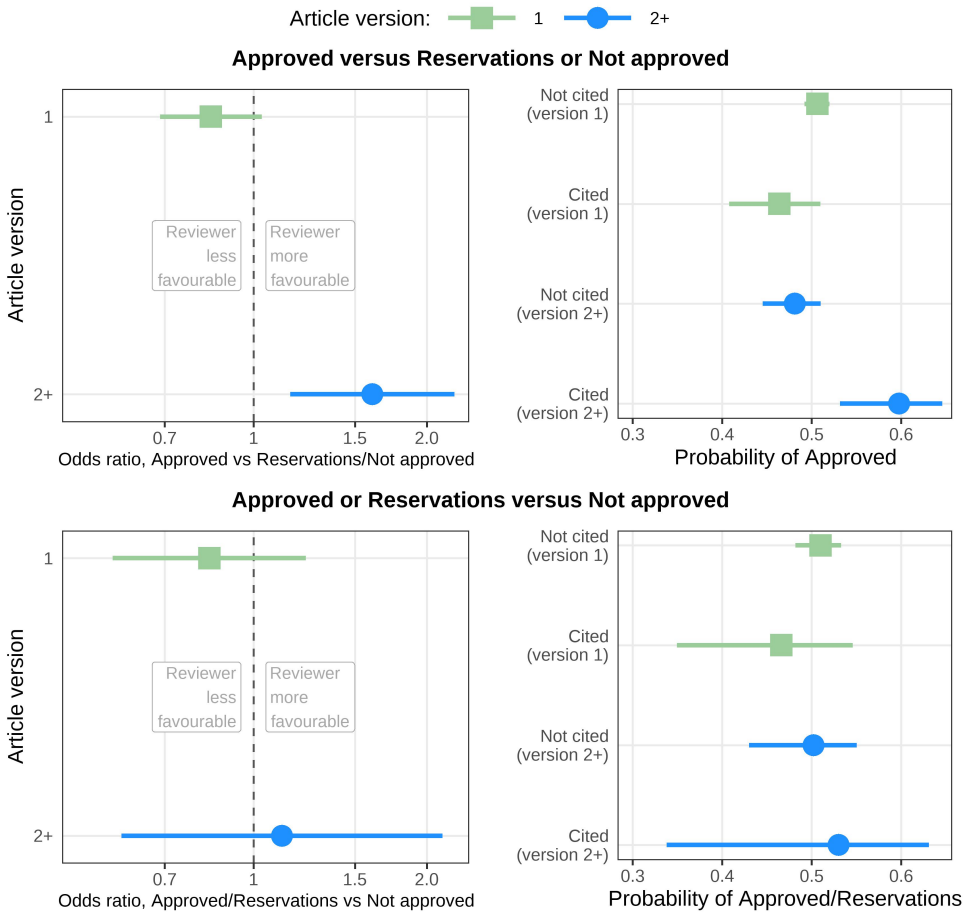
Table 2: **Descriptive statistics for the articles and peer reviews.** *Q1 = first quartile, Q3 = third quartile.*

Variable	Level / Statistics	Result
Number of reviews	n	37,332
Year	Median [Q1, Q3]	2022 [2019, 2024]
Journal, n (%)	F1000Research	24,132 (65)
	Wellcome Open Research	8697 (23)
	Open Research Europe	2789 (7)
	Gates Open Research	1714 (5)
Role, n (%)	Reviewer	34,904 (93)
	Co-reviewer	2428 (7)
Reviewer's recommendation, n (%)	Approved	19,984 (54)
	Reservations	14,379 (38)
	Not approved	2969 (8)
Article version, n (%)	1	26,474 (71)
	2	8995 (24)
	3+	1863 (5)
Number of papers cited in article	Median [Q1, Q3]	24 [14, 38]
Any citations to reviewer, n (%)	No	32,375 (87)
	Yes	4957 (13)
Any papers cited by reviewer, n (%)	No	31,546 (84)
	Yes	5786 (16)
Any citations to the reviewer's articles	No	35,023 (94)
	Yes	2309 (6)
Reviewer's publication count	Median [Q1, Q3]	55 [24, 118]
Reviewer's country (top five only)	USA	7655 (21%)
	United Kingdom	4137 (11%)
	India	2472 (7%)
	Italy	1368 (4%)
	Australia	1349 (4%)
Number of words in the review	Median [Q1, Q3]	202 [67, 411]

The binary predictor for citations of "any versus none" had a generally better fit to the data compared to the linear predictor (Supplement S.3). This indicates that for most reviewers receiving any citation is important, and there is no linear increase for two or more citations. The following results are for the binary predictor "any versus none", with the results using a linear predictor in Supplement S.4.

Reviewers who were cited were more likely to approve the article, but only after version 1 (Fig 2 and Table 3). If a reviewer was cited in any versions after version 1, the odds ratio for

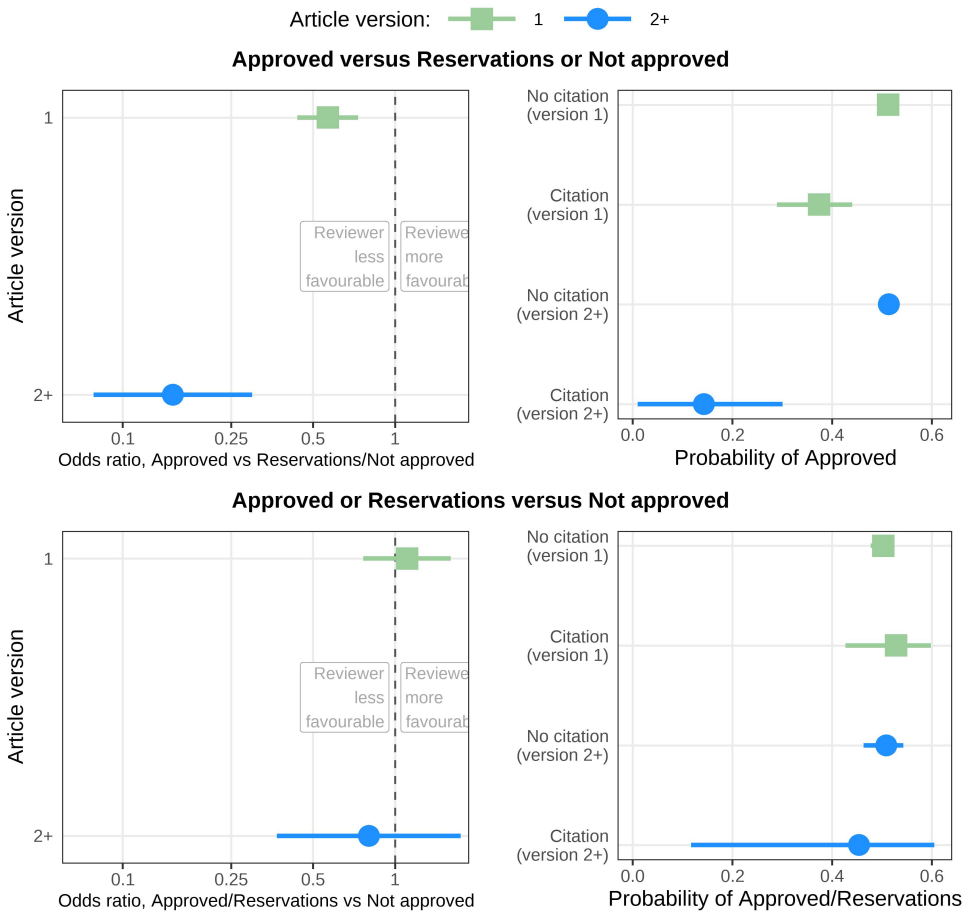
Figure 2: **Odds ratios and probabilities for reviewers giving a more or less favourable recommendation depending on whether they were cited in the article.** *Top left: Odds ratios for reviewers giving a more favourable (Approved) or less favourable (Reservations or Not approved) recommendation depending on whether they were cited in the article. Reviewers cited in later versions (blue) were more likely to make a favourable recommendation (odds ratio = 1.61; adjusted 99.4% CI: 1.16 to 2.23), whereas being cited in the first version (green) did not improve their recommendation (odds ratio = 0.84; adjusted 99.4% CI: 0.69 to 1.03). Top right: Same results as top left displayed as conditional probabilities. From the top, the lines show the within-strata probability of a reviewer approving: a version 1 article in which they are not cited (0.51; adjusted 99.4% CI: 0.49 to 0.52); a version 1 article in which they are cited (0.46; adjusted 99.4% CI: 0.41 to 0.51); a version 2 (or higher) article in which they are not cited (0.48; adjusted 99.4% CI: 0.45 to 0.51); and a version 2 (or higher) article in which they are cited (0.60; adjusted 99.4% CI: 0.53 to 0.65). Bottom left: Same estimates as top left except that a more favourable recommendation is now Approved or Reservations and a less favourable is Not approved. There was no clear association for cited reviewers in version 1 (odds ratio = 0.84; adjusted 99.4% CI: 0.57 to 1.23) or later versions (odds ratio = 1.12; adjusted 99.4% CI: 0.59 to 2.13). Bottom right: Same results as bottom left displayed as conditional probabilities. From the top, the lines show the within-strata probability of a reviewer approving: a version 1 article in which they are not cited (0.51; adjusted 99.4% CI: 0.48 to 0.53); a version 1 article in which they are cited (0.47; adjusted 99.4% CI: 0.35 to 0.55); a version 2 (or higher) article in which they are not cited (0.50; adjusted 99.4% CI: 0.43 to 0.55); and a version 2 (or higher) article in which they are cited (0.53; adjusted 99.4% CI: 0.34 to 0.63). This figure is based on an analysis of 12,051 articles and 24,677 reviews for version 1 and 6090 articles and 10,196 reviews for version 2+. In all panels a dot or square represents a mean, and a horizontal line represents an adjusted 99.4% confidence interval.*



recommending Approved versus Reservations or Not approved was 1.61 (adjusted 99.4% CI 1.16 to 2.23).

Reviewers who included a citation to their own articles were much less likely to approve the article for all versions (Fig 3 and Table 3). The odds ratio for recommending Approved versus Reservations or Not approved was 0.57 (99.4% CI 0.44 to 0.73) for version 1 and strengthened to

Figure 3: **Odds ratios and probabilities for reviewers giving a more or less favourable recommendation if they included a citation to their own articles in their review.** *Top left: Odds ratios for reviewers giving a more favourable (Approved) or less favourable (Reservations or Not approved) recommendation depending on whether their review included a citation to their own articles. Reviewers including a citation to their own articles were less likely to make a favourable recommendation for version 1 (green; odds ratio = 0.57; adjusted 99.4% CI: 0.44 to 0.73) and later versions (blue; odds ratio = 0.15; adjusted 99.4% CI: 0.07 to 0.30). Top right: Same results as top left displayed as conditional probabilities. From the top, the lines show the within-strata probability of a reviewer approving: a version 1 article in which their review did not include a citation (0.51; adjusted 99.4% CI: 0.50 to 0.53); a version 1 article in which in which their review included a citation (0.37; adjusted 99.4% CI: 0.29 to 0.44); a version 2 (or higher) article in which their review did not include a citation (0.51; adjusted 99.4% CI: 0.49 to 0.53); and a version 2 (or higher) article in which in which their review included a citation (0.14; adjusted 99.4% CI: 0.01 to 0.30). Bottom left: Same estimates as top left except that a more favourable recommendation is now Approved or Reservations and a less favourable is Not approved. There was no clear association for reviewers who included a citation to their own articles in version 1 (odds ratio = 1.11; adjusted 99.4% CI: 0.77 to 1.60) or later versions (odds ratio = 0.80; adjusted 99.4% CI: 0.37 to 1.74). Bottom right: Same results as bottom left displayed as conditional probabilities. From the top, the lines show the within-strata probability of a reviewer approving: a version 1 article in which their review did not include a citation (0.50; adjusted 99.4% CI: 0.48 to 0.52); a version 1 article in which their review included a citation (0.53; adjusted 99.4% CI: 0.43 to 0.60); a version 2 (or higher) article in which their review did not include a citation (0.51; adjusted 99.4% CI: 0.46 to 0.54); and a version 2 (or higher) article in which their included a citation (0.45; adjusted 99.4% CI: 0.12 to 0.61). This figure is based on an analysis of 12,078 articles and 24,732 reviews for version 1 and 6101 articles and 10,213 reviews for version 2+. In all panels a dot or square represents a mean, and a horizontal line represents an adjusted 99.4% confidence interval.*



0.15 (99.4% CI 0.08 to 0.30) for versions 2+. The less favourable recommendation was only for the approval of the article and the odds ratios for Approved or Reservations versus Not approved were much closer to 1.

In an unplanned analysis, we examined the behaviour of reviewers in the first two versions of the

Table 3: Odds ratios for reviewers giving a more (OR > 1) or less (OR < 1) favourable recommendation depending on whether they were cited in the article (question 1) or included citations to their own articles (question 2). All models were split by article version.

Research question	Article version	Outcome	OR (Adjusted 99.4% CI)
1. Reviewer cited by authors	Version = 1	Approved vs Reservations/Not approved	0.84 (0.69, 1.03)
	Version = 1	Approved/Reservations vs Not approved	0.84 (0.57, 1.23)
	Versions = 2+	Approved vs Reservations/Not approved	1.61 (1.16, 2.23)
	Versions = 2+	Approved/Reservations vs Not approved	1.12 (0.59, 2.13)
2. Reviewer cited their own articles	Version = 1	Approved vs Reservations/Not approved	0.57 (0.44, 0.73)
	Version = 1	Approved/Reservations vs Not approved	1.11 (0.77, 1.60)
	Versions = 2+	Approved vs Reservations/Not approved	0.15 (0.08, 0.30)
	Versions = 2+	Approved/Reservations vs Not approved	0.80 (0.37, 1.74)

article. We examined the 441 reviews where the reviewer was not cited in version 1 of the article and included a citation to their own articles in their first review. The reviewers who were then cited in version 2 recommended approval for 92% compared to 76% for reviewers who were not cited (odds ratio = 3.5, 95% CI: 2.0 to 6.1). This analysis did not use matching.

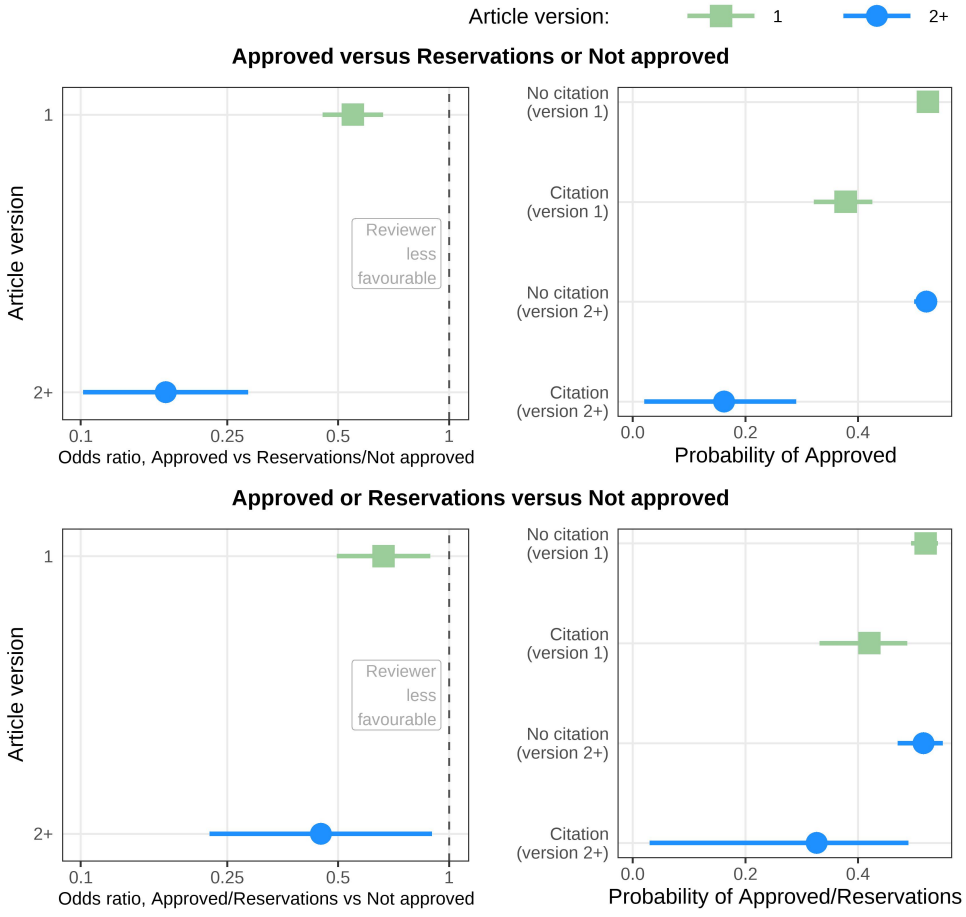
In an unplanned analysis, we examined whether the reviewers' recommendations depended on whether their review included citations to articles other than their own. Reviewers who included citations in their review were much more likely not to approve the article (Figure 4), which was similar to the association with citations to reviewers' own articles (Figure 3). However, reviewers who included citations to articles other than their own were also much more likely to recommend "Not approved", as shown by the lower odds of "Approved" or "Reservations" versus "Not approved". This association was not seen using citations to reviewers' own articles (Figure 3)

Sensitivity analyses

The odds ratios when including co-reviewers with reviewers were similar to the odds ratios when using reviewers only (Supplement S.5).

We found no evidence that the reviewers' publication numbers or country confounded the associations between citations and recommendations (Supplement S.6).

Figure 4: **Odds ratios and probabilities for reviewers giving a more or less favourable recommendation depending on if they included citations to articles other than their own in their review.** *Top left: Odds ratios for reviewers giving a more favourable (Approved) or less favourable (Reservations or Not approved) recommendation depending on whether their review included a citation to articles other than their own. Reviewers including citations to other articles were less likely to make a favourable recommendation for version 1 (green; odds ratio = 0.53; adjusted 99.4% CI: 0.44 to 0.64) and later versions (blue; odds ratio = 0.18; adjusted 99.4% CI: 0.10 to 0.30). Top right: Same results as top left displayed as conditional probabilities. From the top, the lines show the within-strata probability of a reviewer approving: a version 1 article in which their review did not cite other articles (0.53; adjusted 99.4% CI: 0.51 to 0.54); a version 1 article in which their review cited other articles (0.37; adjusted 99.4% CI: 0.31 to 0.42); a version 2 (or higher) article in which their review did not cite other articles (0.52; adjusted 99.4% CI: 0.50 to 0.54); and a version 2 (or higher) article in which in which their review cited other articles (0.17; adjusted 99.4% CI: 0.02 to 0.30). Bottom left: Same estimates as top left except that a more favourable recommendation is now Approved or Reservations and a less favourable is Not approved. Reviewers including citations to other articles were less likely to make a favourable recommendation for version 1 (odds ratio = 0.62; adjusted 99.4% CI: 0.46 to 0.84) and later versions (odds ratio = 0.34; adjusted 99.4% CI: 0.16 to 0.73). Bottom right: Same results as bottom left displayed as conditional probabilities. From the top, the lines show the within-strata probability of a reviewer approving: a version 1 article in which their review did not cite other articles (0.52; adjusted 99.4% CI: 0.49 to 0.54); a version 1 article in which their review cited other articles (0.41; adjusted 99.4% CI: 0.31 to 0.48); a version 2 (or higher) article in which their review did not cite other articles (0.52; adjusted 99.4% CI: 0.47 to 0.55); and a version 2 (or higher) article in which their review cited other articles (0.27; adjusted 99.4% CI: 0.02 to 0.45). This figure is based on an analysis of 12,078 articles and 24,732 reviews for version 1 and 6101 articles and 10,213 reviews for version 2+. In all panels a dot or square represents a mean, and a horizontal line represents an adjusted 99.4% confidence interval.*

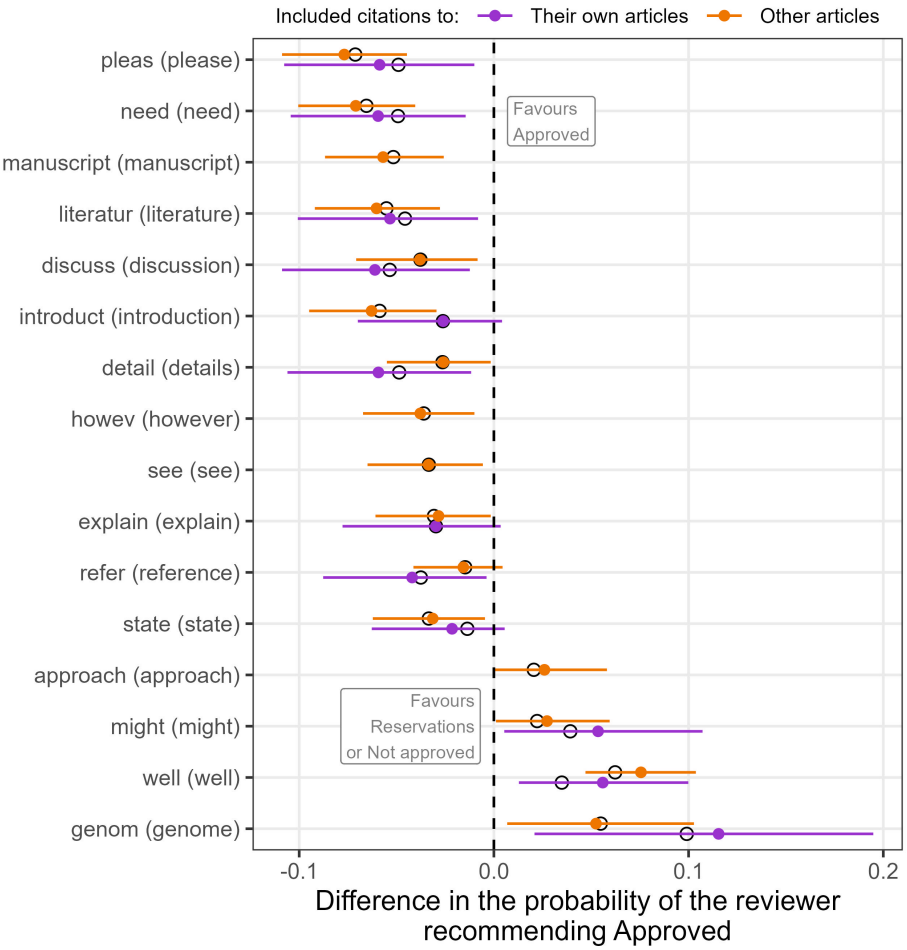


Text analyses of reviewers' comments

A random sample of how reviewers included citations to their own articles found some vague justifications (Supplement S.7); for example, "Here are some additional publications you might consider referencing". Other sentences adhered to the publisher's guidelines for reviewers, as specific reasoning was provided for citations to their own articles [7]. One reviewer thanked the authors for a previous citation. Three reviews did not have a relevant sentence. One reviewer almost certainly used AI to write their review as it included the phrase "Certainly! Here are some potential review questions for the manuscript" [43]; this review included six self-citations with no justifications.

Reviewers who included a citation to their own articles or other articles were more likely to use the words "need" and "please" when not approving the article (Figure 5). In contrast, the words "genome" and "well" were the most strongly associated with the reviewers' approval.

Figure 5: **Words in the reviewers' comments that were associated with approving the article or not for reviewers who included a citation to their own articles ($n = 2,025$) and reviewers who included citations to other articles ($n = 4,350$).** *The words were selected using an elastic net that started with the 100 most commonly used review words. The estimates from the elastic net are shown as empty circles and the mean estimates and 95% credible intervals from a Bayesian model are shown as a solid circle and horizontal line. Words are shown if the probability of a non-zero mean was over 0.95 for either reviewers who cited their own articles or reviewers who cited other articles. Four words were selected by the elastic net for the reviewers who cited other articles but not by the elastic net for reviewers who cited their own articles. The axis label shows the stemmed word and most common whole word in brackets.*



To examine how often open peer reviews were viewed, we took a random sample of 200 reviews from the four journals and found that, on average, they were viewed just 1.2 times per year (Supplement S.8).

Discussion

Our results provide evidence that some reviewers have a transactional view of peer review, with their final approval dependent on citations to their work. Some reviewers may be exploiting the pressure on authors to "publish or perish". Under this pressure, many authors may oblige and add the suggested citations, especially since adding another citation may only require a minor edit to their article [44]. Both sides gain from this transaction, as the authors get an indexed publication and the reviewer gets a citation.

A key question is whether citations to a reviewer's own articles are justified as they may highlight important errors or missing context in the article. Citations to a reviewer's own articles can be justified when the authors have made a "large scholarly oversight" [8]. To investigate this, we compared the recommendations and wording of reviewers who included citations to their own articles to reviewers who included citations to other articles. The language used by the reviewers of these two groups was similar, with a higher use of "please" and "need" when not approving the article (Figure 5). However, there was a difference between groups in their recommendations, as reviewers including citations to other articles were more likely to recommend "Not approved" (Figure 4) whereas this association was not observed for reviewers including citations to their own articles (Figure 3). This indicates that missing citations to other articles were considered more serious than missing citations to the reviewer's articles. Reviewers who cited their own articles may have been more inclined to give authors a chance to update their article and thus potentially include the "missing" citation(s).

Examining the context of the citations to reviewers' own articles, we found vague or non-existent justifications (Supplement S.7), showing that some reviewers ignored the journals' guidelines to state their reasoning when including citations to their own articles. However, these examples of poor justifications do not mean that all self-citations are coercive.

For both research questions, the effects were stronger for the second and later versions of the article than for the first version. Reviewers may understand that authors may be more willing to compromise on later versions when they are closer to obtaining an indexed publication. Most researchers understand that the peer review system is imperfect and that they sometimes have to make compromises to be successful [20, 45]. Another difference to consider is that later versions will include more articles with disagreements between reviewers and more that were not "Approved" in the first version as articles where two or more reviewers recommended "Approved" may not have needed a second version.

Potential improvements to peer review

Journals could give stronger guidance to reviewers and authors on coercive citations [4]. However, given the limited time for peer review and the many differences in guidelines between journals [46], most authors may not read peer review instructions. Hence, guidance alone may have limited impact.

One suggestion is that reviewers declare to editors when they have recommended citations to their own work [47]. A useful innovation would be for all reviews that contain citations to the reviewer's own articles to be automatically flagged to the editors who could check if the citations are justified. We are aware of one journal where this is already happening (personal communication, Benno Torgler). *F1000* have recently introduced checks to prevent reviewers from publishing a review with three or more citations to the reviewer's own articles. If the reviewers continue to request more than three, then the review is examined, and if the citations are deemed inappropriate and the reviewer declines to remove them, then the review is declined.

Open peer review has been suggested as a way to reduce coercive citations [7, 47]. However, our results from four journals that use open review show that it is not a perfect antidote, although the problem could be worse in journals using blinded peer review. The transparency of open peer review should prevent reviewers from leaving self-serving comments; however, we found some dubious justifications for self-citations and blatant use of AI (Supplement S.7). These reviewers may have rationalised that although their words are public, they are rarely scrutinised (Supplement S.8); hence, it was worth the risk. The assumed additional quality assurance from open peer review [48] may often be absent.

A more radical change to peer review is that the reviewers initially see a version of the article with all references blinded and no reference list; for example, "A strong association between increased cleaning and reduced hospital infection is well established [x]". Reviewers are asked to give an initial recommendation and comments, and then are shown the version with the full references and asked if they need to update their recommendation or provide additional comments. However, this involves more administrative work and demands more from peer reviewers. This approach could be used for particularly consequential or controversial articles. Some journals already require authors to partially blind their articles to maintain anonymous peer review; for example, the instructions from *Taylor & Francis* include blinding the authors' names in the reference list [49].

An argument could be made for using large language models to provide peer review that is unmoved by citation flattery. However, peer review is an inherently human task by peers, and instead we need to improve peer review rather than abdicating this often difficult and time-consuming task to machines [50].

Related research

Previous cross-sectional studies of self-citations in reviews found at least one self-citation in 3% at a journal that used blinded peer review [17], 12% at a journal that used blinded peer review [51], and 12% at a journal that used open peer review [52]. A related study found that 15% of reviews included a self-citation and that the self-citations were highest when the reviewer recommended "major revisions" [53]. These figures are comparable with the 6% found here and indicate that most reviews do not include self-citations.

Previous surveys estimated that 14% and 20% of authors had experienced a coercive citation request from an editor [54, 55], and 7% and 23% had experienced coercive citation pressure from a reviewer [14, 56]. The frequency with which researchers interact with peer review means that many will encounter coercive citations at some point in their career.

A study of conference submissions estimated that reviewers who were cited gave submissions much higher scores [18]. A study of journal peer review estimated that cited reviewers scored the article higher, but with potential confounding by the quality of the article [17].

A survey of authors concluded that accepting an editor's request for citations improved the chances of being accepted [12]. Requests in later versions were more strongly associated with acquiescence, and we found a related pattern in our analysis, with reviewers who included citations to their own articles being much less likely to recommend approval for later versions (Figure 3). A study examining open peer review found that requests to cite the reviewer's articles were more likely to be included than other suggested citations, indicating that many authors wanted to please the reviewer or felt pressure to do so [52].

A survey of journal editors found that only 5% objected to reviewers citing their own articles, and that this should be expected as reviewers are likely to have done related work [8].

A cross-sectional study found that reviewers citations to their own articles were more likely to have no rationale compared to other citations, suggesting that they are more likely to be unwarranted [51].

Strengths and limitations

This is an observational study meaning we cannot rule out unmeasured confounding and should be cautious in interpreting the results.

To our knowledge, this is the first analysis to use a matched design and analysis when examining reviewer citations, and hence strongly control for any confounding by the characteristics of the authors or articles. We compared reviewers who examined an identical article; hence, the differences we found should be due to the reviewers.

Our models include measurement error, as some citations to the reviewers' work will be missed by our data collection, and some captured citations will be inaccurate [57]. We performed random data checks that showed good accuracy (Supplement S.9); however, we also found valid citations that were not captured by our data extraction for conference proceedings and technical

reports, which are less likely to have a DOI. This measurement error would most likely underestimate a true association, as it reduced the variance in citation counts and created a regression dilution [58]. Our estimates will be biased if the associations between citations and reviewers' recommendations are different for publications that do not have a DOI. Reviewers should be equally happy with any citation to their work; however, some reviewers may prefer citations to indexed articles, as these are more likely to count toward their h-indices [59].

We examined whether citing a reviewer altered their recommendation, but did not examine the sentiment of the citation [60]. Some citations would likely have been critical of the reviewer's articles, and we would expect these to reduce the chances of a favourable recommendation. An analysis that included the sentiment of the citation would be useful, although previous research found that most citations are neutral or positive [60].

We did not examine the authors' responses to the reviewers but these could include important information on why a citation was included or not in a revised version of the article. A detailed analysis examining the text used in the interactions between authors and reviewers could provide valuable information about the peer review process.

Our results may not be generalisable to journals that use blinded peer review or journals that use the traditional peer review model rather than the publish–review–curate model studied here. A previous study found that asking reviewers to consent to an open review had no important effect on the quality of the review or the reviewers' recommendation [61]. Another potential difference is that the journals in our sample often asked the authors to suggest peer reviewers; however, this is relatively common in other journals [8].

We found a bias in our sample, as co-reviewers and reviewers from older articles were more likely to be excluded due to not having an *OpenAlex* record (Supplement S.6). We therefore lost more junior reviewers who were less likely to be cited. The percentage of reviews lost was 5% (2,026 of 39,113), which is hopefully small enough to avoid a large bias.

Acknowledgements

Thanks to all four journals for making all their data openly available and easily accessible. Thanks to Robin Blythe, staff from *F1000* and Paper-Wizard `https://paper-wizard.com/` for providing helpful feedback on a draft of this paper.

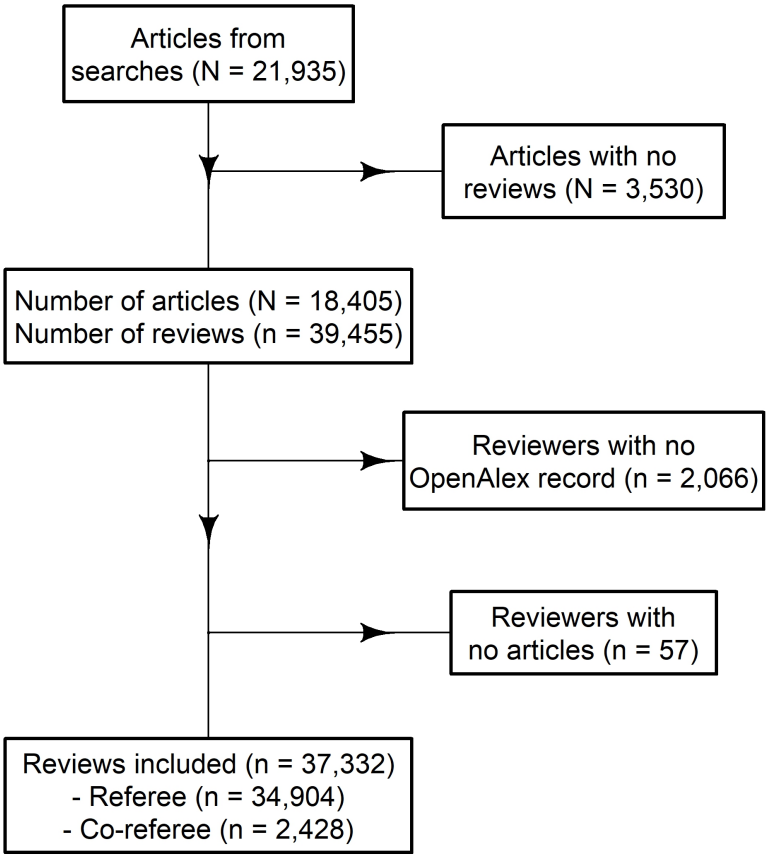
Supplementary material

S.1 Sample size

We aimed for a sample size of approximately 5,000 articles and assumed that half would be the first version, giving a sample size of 2,500 articles for the analysis using the first version only [40]. In 1,000 simulations, this gave an 89.1% power to detect an odds ratio of 1.5 using conditional logistic regression for a reviewer who recommended a higher category (Approved → Reservations → Not approved) when they were cited. We assumed that 15% of articles would include a citation to the reviewer. Eighty percent of the simulated articles had two reviews, and the remaining 20% had three reviews. Based on preliminary data from two journals, we assumed that the reviewers' recommendations would have a ratio for Approved:Reservations:Not approved of 70:24:6.

S.2 Included and excluded reviews

Figure S.1: **Flow chart of included reviews. 'N' is the number of articles and 'n' is the number of reviews.**



The flow chart shows the loss of articles and reviews during the data collection process. More than 3,500 articles did not have reviewers as they had yet to be peer reviewed or were Faculty Reviews that are commissioned and use a different peer review model.

More than 2,000 reviewers did not have an *OpenAlex* record and therefore were excluded from the analyses. We examined the potential bias in the lost reviews by comparing their characteristics with those of the retained reviews. We used a multiple regression model with reviewer lost (yes/no) as the binary dependent variable and predictors of article version, article date, referee or co-referee, and reviewer's country. We expected many of these predictors to have little effect; therefore, we used an elastic net to reduce the number of predictors [38]. We used the 'glmnet' package in *R* [39]. For the binary dependent variable, 39,455 reviews were retained and 2082 (5%) were lost.

The elastic net retained two predictors. The date of the article had an odds ratio of 1.09 per year increase, which means that more recent articles were more likely to be retained, likely because the reviewer's information was more current. Referees were more likely to be retained compared to co-referees with an odds ratio of 1.79, likely because co-referees were often relatively junior and some may not have any publications.

S.3 Model fit

Table S.1: **Comparing the two alternatives for the citation predictor variables using either a linear variable or a binary "any versus none" variable.** *A vs R/N = Approved vs Reservations/Not approved, A/R vs N = Approved/Reservations vs Not approved.*

Co-reviewers included	Version	Outcome	AIC		
			Linear	Binary	Difference
No	1	A vs R/N	5940.9	5937.4	3.6
No	1	A/R vs N	1930.3	1929.8	0.5
No	2+	A vs R/N	1952.4	1941.9	10.5
No	2+	A/R vs N	572.1	571.9	0.2
Yes	1	A vs R/N	5978.4	5975.4	3.0
Yes	1	A/R vs N	1941.1	1940.8	0.3
Yes	2+	A vs R/N	1963.1	1951.4	11.7
Yes	2+	A/R vs N	572.6	572.8	–0.2
No	1	A vs R/N	5932.3	5911.1	21.2
No	1	A/R vs N	1934.1	1935.6	–1.5
No	2+	A vs R/N	1881.4	1876.0	5.4
No	2+	A/R vs N	573.4	572.8	0.6
Yes	1	A vs R/N	5967.9	5944.9	23.0
Yes	1	A/R vs N	1945.4	1946.6	–1.2
Yes	2+	A vs R/N	1917.3	1904.1	13.2
Yes	2+	A/R vs N	573.1	573.5	–0.5

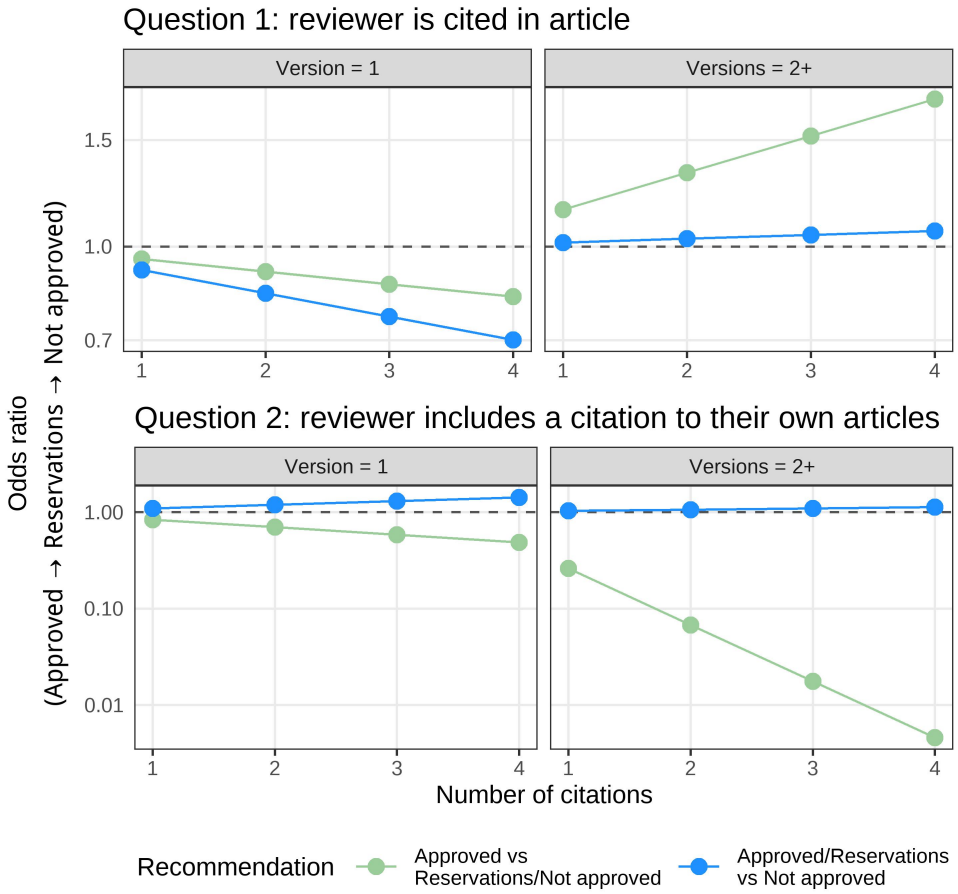
The AIC (Akaike Information Criterion) is a trade-off of model fit and complexity. The smaller the AIC, the better the fit. Differences of 10 are considered large [35].

In most cases, the difference between the linear and binary variables was small (under 5). There were four comparisons out of 16 in which the linear variable had a smaller AIC than the binary variable and all differences were small (under 2). There were four comparisons where the AIC for the binary variable was over 10 units smaller than the linear variable, indicating a large difference in model fit. In summary, using a binary predictor variable is a generally better fit to the data than using a linear variable.

S.4 Results from using a linear predictor

The figure shows the estimates for the two research questions using a linear dose-response for citation counts instead of the binary predictor of any citation versus none. The strongest effect was a greatly reduced odds of "Approved" for increasing citations to the reviewer's own articles. However, these estimates should be viewed with caution, as the binary predictor generally better fits the data (Supplement S.3).

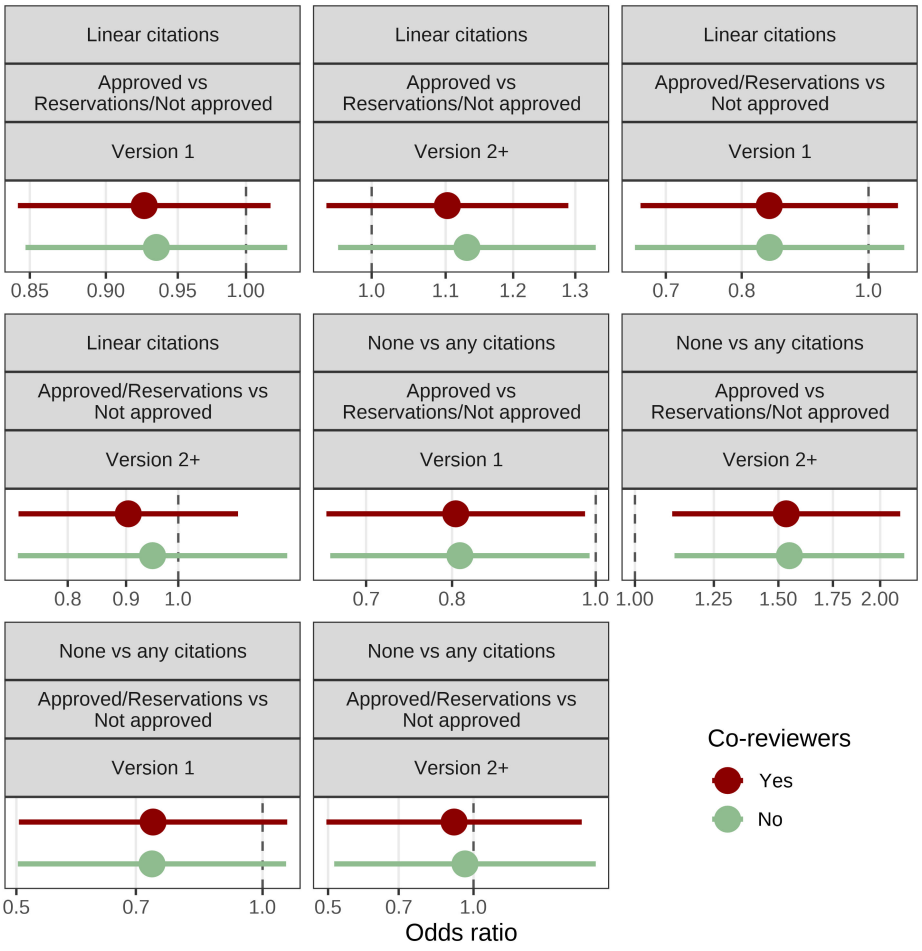
Figure S.2: **Estimated odds ratios for using linear citations as the predictor.** *The reference point is zero citations.*



S.5 Including co-reviewers

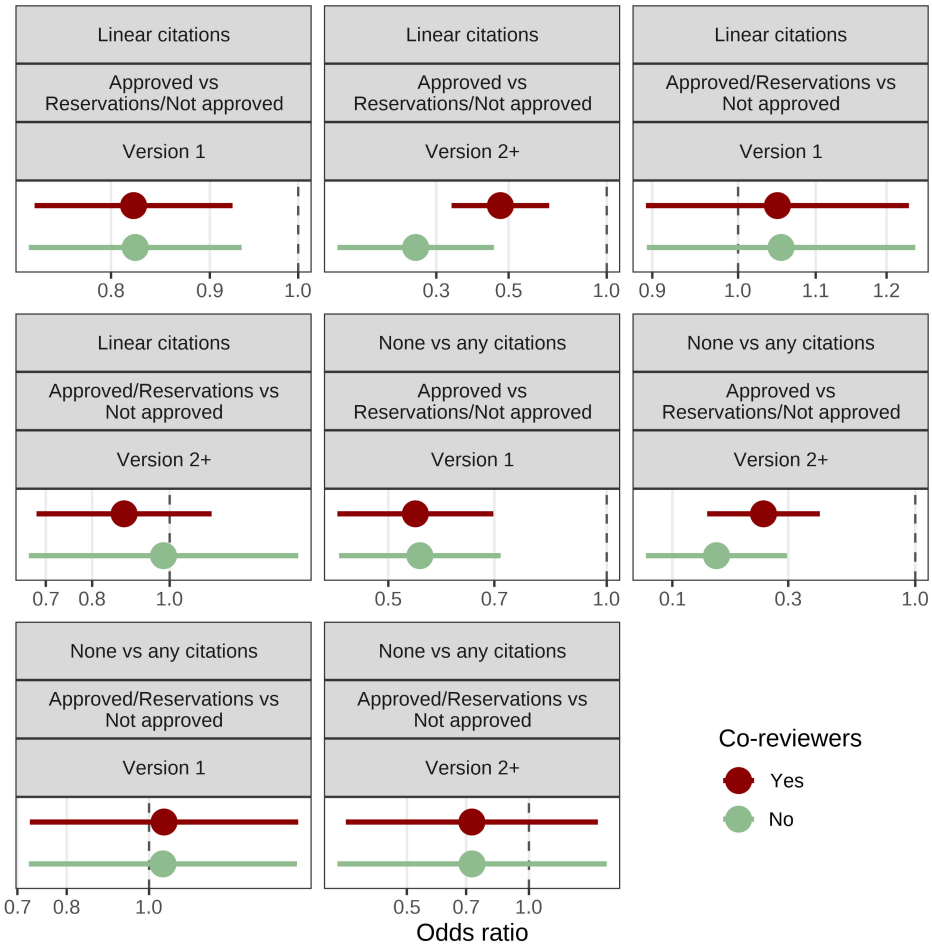
Some reviews were performed by reviewers together with co-reviewers, who were usually less experienced. Our primary analysis excluded co-reviewers, but we included them in a sensitivity analysis where we created combined versions of the two independent variables using the sum of citations to reviewers and co-reviewers, and the sum of citations to their own articles from the reviewers and co-reviewers. The results examining whether the reviewers gave a more favourable recommendation when cited (research question 1) were very similar (Figure S.3).

Figure S.3: **Results with or without co-reviewers for research question 1.** *Odds ratios and adjusted 99.4% confidence intervals for whether the reviewer gave a more or less favourable recommendation if they were cited. The results are shown for the combinations of predictor variables (linear or any vs none), outcome (Approved → Reservations → Not approved) and article version. The plot is designed to directly compare paired odds ratios with or without co-reviewers.*



The results examining whether the reviewers gave a more favourable recommendation when they included citations to their own articles (research question 2) were mostly very similar (Figure S.4). Two noticeable differences were two odds ratios where including co-reviewers somewhat reduced the strength of the association. This was for article versions 2+ and examining Approved vs Reservations or Not approved. Despite the noticeable change in the odds ratio, the interpretation remains similar in that there was a strong reduction in the odds of a favourable recommendation when the reviewers included citations to their own articles.

Figure S.4: **Results with or without co-reviewers for research question 2.** *Odds ratios and adjusted 99.4% confidence intervals for whether the reviewer gave a more or less favourable recommendation when they included a citation to their own articles. The results are shown for the combinations of predictor variables (linear or any vs none), outcome (Approved → Reservations → Not approved) and article version. The plot is designed to directly compare paired odds ratios with or without co-reviewers.*



S.6 Potential confounding by the reviewers' characteristics

Any confounding by the characteristics of the articles was controlled by the matched design, but confounding by the characteristics of the reviewers remains possible [18]. We considered the potential confounders of the reviewer's experience and reviewer's country. More experienced reviewers will likely be cited more often (on average) and could be more or less strict in their recommendations. The reviewer's country is a potential confounder due to large differences in citation counts by country [62] and potential differences in recommendations by country [63].

Reviewers' experience

We used the reviewer's number of published articles as a proxy for their experience. This association could be non-linear; for example, a diminishing effect for more experienced reviewers, so we examined six fractional polynomials of the reviewers' number of articles and used the AIC to select the best fit [64]. For most models, the best fit was achieved using a log-transformation.

There was little evidence of any confounding by the reviewers' publication counts as the odds ratios were similar for both research questions (Figures S.5 and S.6). A fractional polynomial of -2 tended to show the largest difference compared to the odds ratios with no confounders; however, this transformation was not the best fit and the differences were relatively small.

Figure S.5: **Examining potential confounding by reviewers' publication counts for research question 1.** *Odds ratios and adjusted 99.4% confidence intervals for whether the reviewer gave a more or less favourable recommendation when they were cited. We used fractional polynomials to examine a potentially non-linear association between reviewers' publication counts and recommendation. The results for "None" are the results without the potential confounder. The results are shown for the combinations of predictor variables (linear or any vs none), outcome (Approved → Reservations → Not approved) and article version. Results are missing when the model did not converge.*

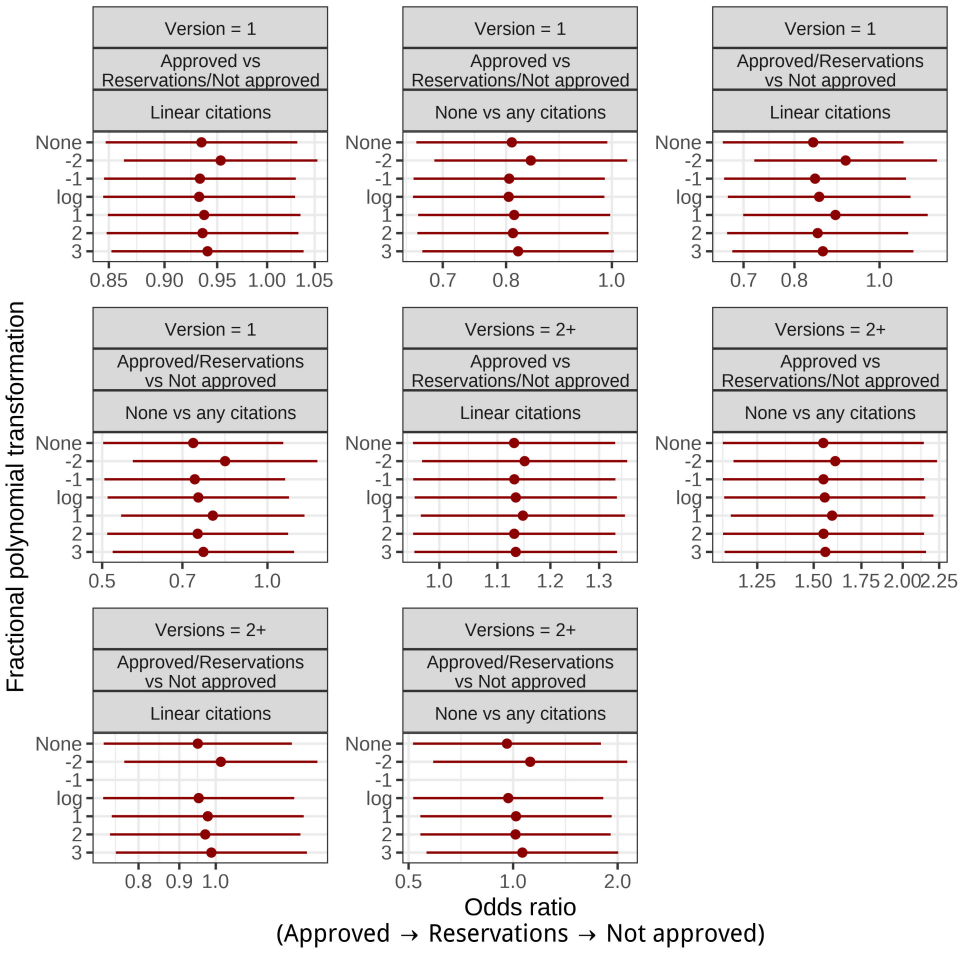
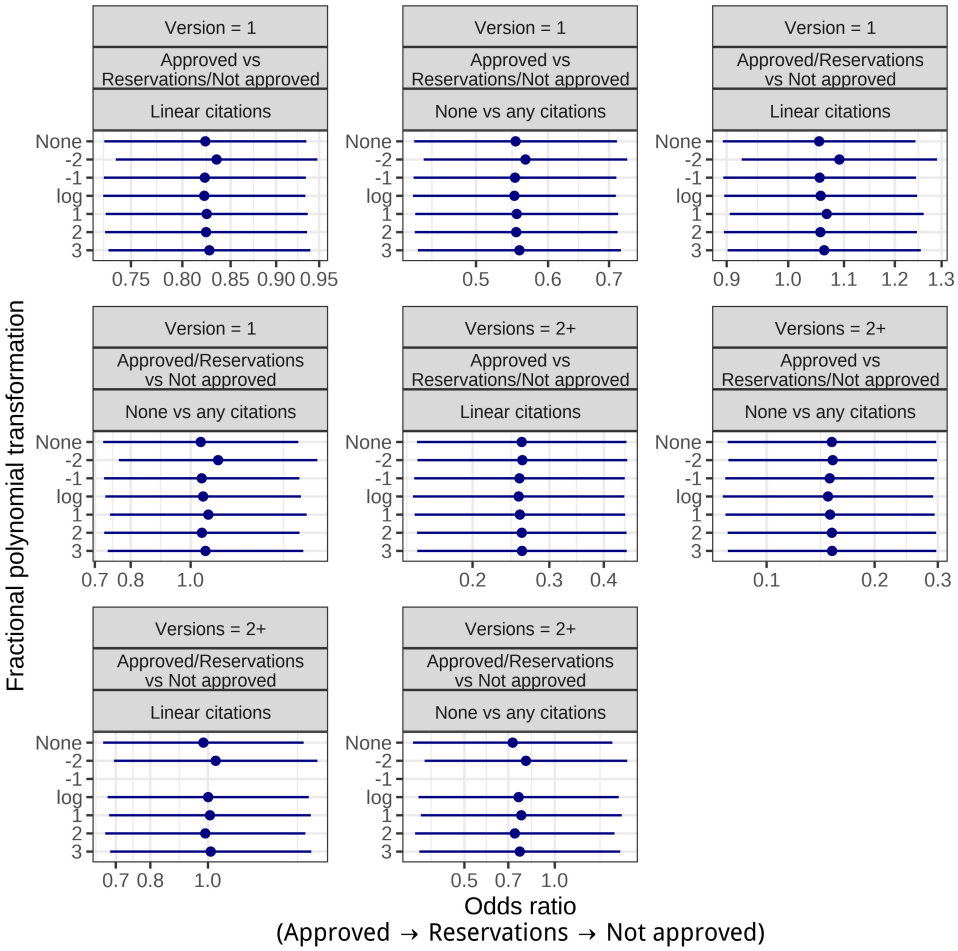


Figure S.6: **Examining potential confounding by reviewers' publication counts for research question 2.** *Odds ratios and adjusted 99.4% confidence intervals for whether the reviewer gave a more or less favourable recommendation when they included a citation to their own articles. We used fractional polynomials to examine a potentially non-linear association between reviewers' publication counts and recommendation. The results for "None" are the results without the potential confounder. The results are shown for the combinations of predictor variables (linear or any vs none), outcome (Approved → Reservations → Not approved) and article version. Results are missing when the model did not converge.*



Reviewers' countries

We planned to use a frailty model to test for confounding by the reviewers' countries [65]. However, this model often failed to converge, potentially because there were many countries and some countries had relatively small numbers of reviewers. Hence, we instead used a leave-one-out analysis for each of the top ten most common countries and determined if the results were noticeably different.

The results were generally similar regardless of which country was left-out. Leaving out the USA, which was the largest country, had a relatively large effect on the odds of recommending Approved or Reservations vs Not approved for versions 2+ when using the "none vs any citations" predictor (Figure S.7) and on the odds of recommending Approved or Reservations vs Not approved for versions 2+ when using the "none vs any citations" predictor (Figure S.8). However, neither change was substantively different from the results including all countries.

Figure S.7: **Leave-one-country-out sensitivity analyses for research question 1.** *Odds ratios and adjusted 99.4% confidence intervals for whether the reviewer gave a more or less favourable recommendation when they were cited. The results are shown for the combinations of predictor variables (linear or any vs none), outcome (Approved → Reservations → Not approved) and article version.*

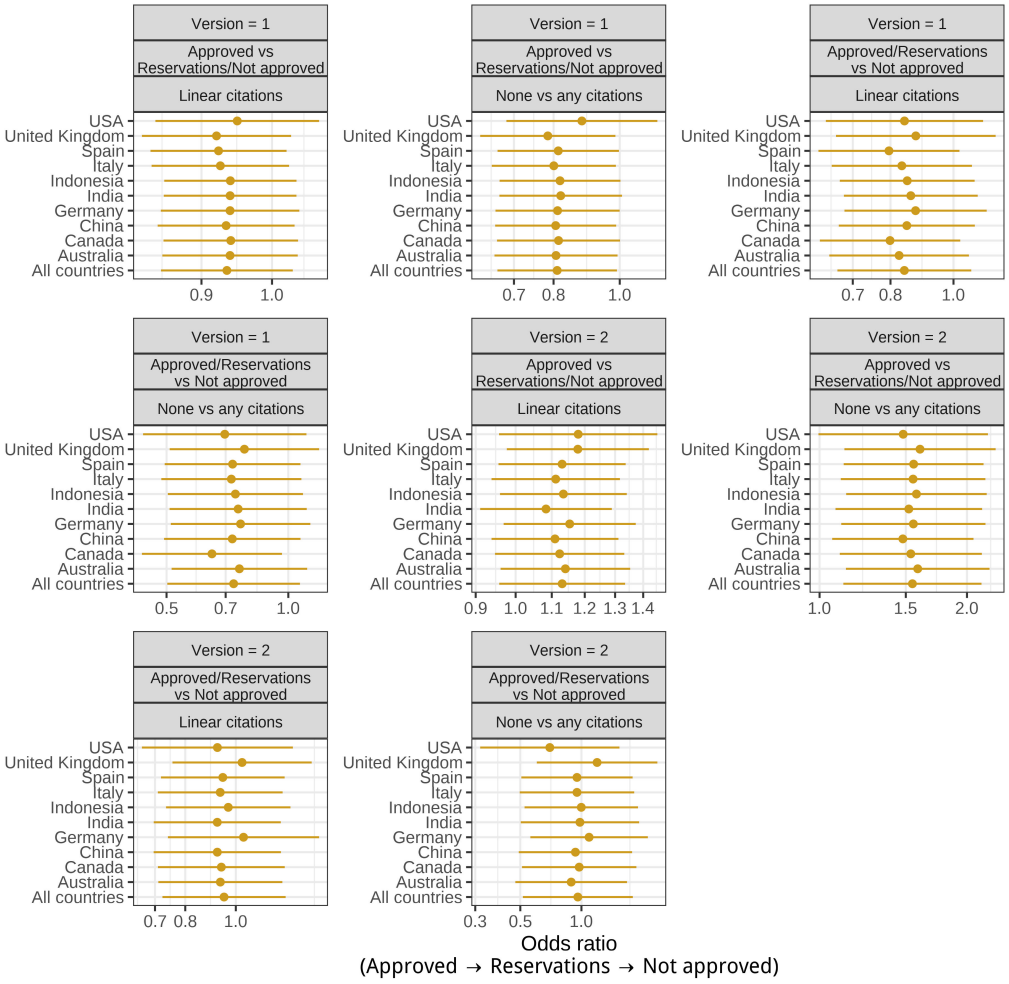
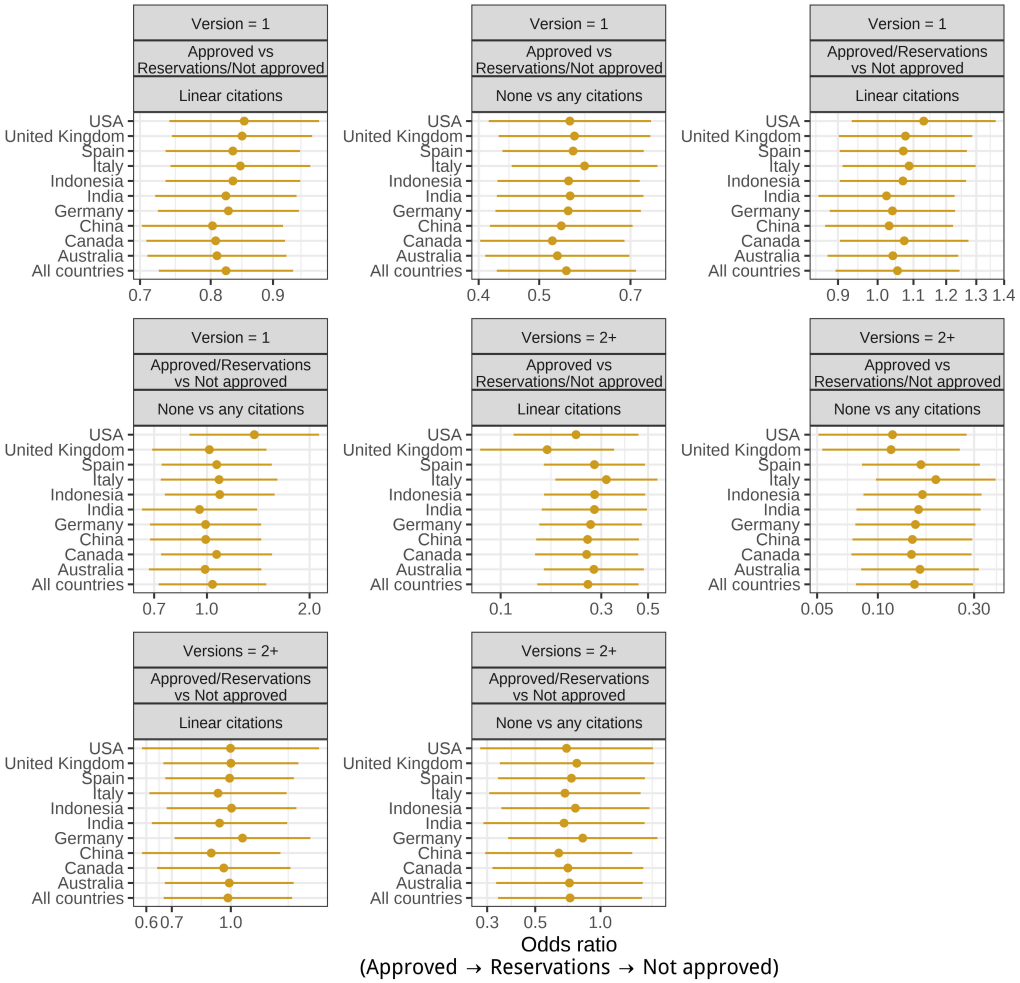


Figure S.8: **Leave-one-country-out sensitivity analyses for research question 2.** *Odds ratios and adjusted 99.4% confidence intervals for whether the reviewer gave a more or less favourable recommendation when they included a citation to their own articles. The results are shown for the combinations of predictor variables (linear or any vs none), outcome (Approved → Reservations → Not approved) and article version.*



S.7 Examples of reviewers' requests to cite their own articles

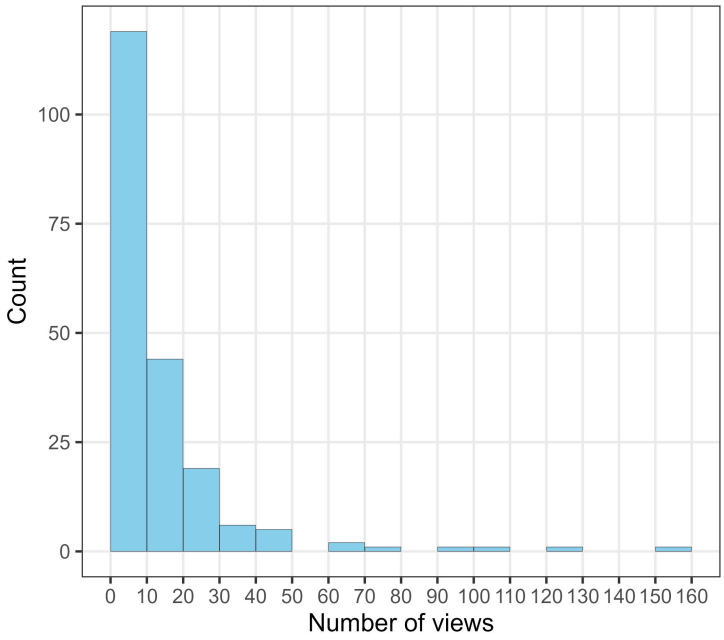
Table S.2: **Example sentences that reviewers used when suggesting citations to their own articles using a random sample of 20 reviews.** *The first column shows the number of citations suggested. We have removed any references to names using [xxxx]. The results are ordered by text length.*

Citations suggested	Reviewer's text
3	Also, the introduction, main discussion, and conclusion must be redrawn to highlight NO as a treatment option, the clinical trials discussion, the use of several NORMS (NORM-1, NORM-2, etc), the effect of NO-carriage system, Natural NO-sources, synthetic NO-sources with limitations, Inorganic versus organic forms, etc (eg., in the review publications as given
1	The term 'true bugs' applies to the monophyletic Heteroptera, which does include the species presented here (Acanthosoma haemorrhoidale), while aphids and mealybugs belong to the distinct lineage Sternorrhyncha, sometimes (formerly) regarded as a part of the paraphyletic Homoptera (see, for example, Figure 2 in: [xxxx]).
1	On a side note: There is already published work on population genomics of the European plaice showing that two large chromosomal rearrangements (two putative inversions) segregate in northern plaice populations (North Sea, Baltic Sea, Barents Sea, and Iceland) and distinguish different plaice populations.
1	There is some observational clinical data on how the detrusor compensates for the growing prostate and the-by consequence-increase in bladder outflow obstruction, in addition to the animal studies referred to in the commentary, to explain the pathophysiology.
1	I would like to thank the authors for including references to the work done in the [xxxx] project; I would recommend to remove the reference ([xxxx] et al., 2019b) and replace it with a reference to a much more recent and related article ([xxxx] et al.):
2	The cited literature is incomplete; it does not include all reports of studies on the presence of the snail in Colombia, and studies with relevant findings of nematodes with or without pathogenic potential in animals are omitted e.g:
1	In the last 2 decades, our group has developed a brief instrument to assess the presence and the severity of sensory phenomena (the University of [xxxx]) to investigate OCD phenotypic subtypes and its relationship with TS/CTD.
5	The authors can find the following relevant articles to enhance their Materials and Methods section and incorporate citations to support their revised manuscript.
3	The authors should consider references from high impact journal publications on crop yield prediction. For example, the following articles by this reviewer
1	No mention to more rigorous rankings such as the Leiden Ranking are made nor to what exactly rankings are portraying. See for instance [xxxx].
1	Maybe 'manifest' and 'not manifest' would work better [for example we used this terminology in [xxxx].
1	This is especially useful if you have multiple data sets - see, for example, the [xxxx] package.
3	I would suggest the authors include some of the results of a large-scale project in Europe.
2	Please refer to some further references to revise the relevant description:
1	Here are a few additional publications you might consider referencing.
1	However, genomics resources are limited, except for parasitoid wasps.
1	Refer to this recent literature review.
4	*No relevant sentence*
1	*No relevant sentence*
6	*No relevant sentence*

S.8 Views of reviews

We randomly sampled 200 reviews from our sample and collected the number of times the review had been viewed online. A histogram of view counts is shown in Figure S.9, which had a strong positive skew with most reviews having 10 or fewer views. We used a Poisson model to estimate the annual number of views per year, accounting for the reviews' publication dates. The mean number of views per year was 1.24 with a 95% credible interval of 1.20 to 1.28.

Figure S.9: **Histogram of online view counts of published reviews. The bins are in tens starting at [0, 10).**



S.9 Data validation

We randomly selected reviews from our analysis data and manually verified the accuracy of our automated data extraction. We checked the accuracy of:

- Reviewers that were cited
- Reviewers that were not cited
- Reviewers that included citation(s) to their own articles in their review

We used a Bayesian calculation to estimate the error rates of our data extraction. We started with a vaguely informative Beta$(1, 3.32)$ prior, which had a 90% probability that the error rate was under 0.5. This vague prior was used to exclude high error rates which were unlikely given our testing of the code during the construction of the data extraction. We created posterior estimates for the error rates using the observed counts of errors from manual checks. We calculated the 90% limits for the posterior distributions as an upper estimate of the error rates.

The distributions are plotted in Figure S.10 and the error rates are shown in Table S.3. The errors are proportions, with 0 for no errors and 1 for all errors. The highest error rate was for citation(s) to their own articles.

Figure S.10: **Distributions of the error rates. Vaguely informative prior and posteriors for errors for not cited reviewers, cited reviewers and self-citations.** *The dashed vertical lines are at* $Pr(error \leq x) = 90\%$.

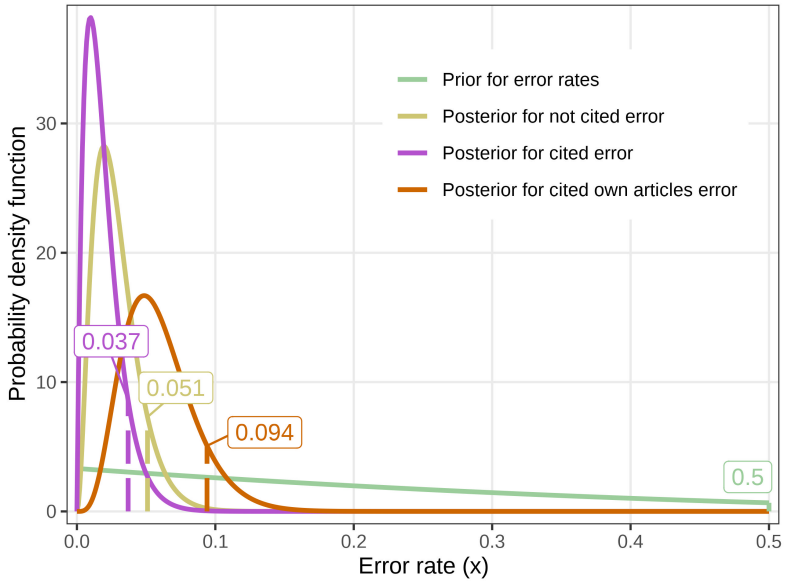


Table S.3: Number of errors found in our data extraction algorithm from manual checks and the estimated 90% limit for the error rate

Check	Number checked	Errors found	Pr(Error rate $\leq x$) = 90%
Reviewer not cited	100	2	0.051
Reviewer cited	100	1	0.037
Reviewer's citation to their own articles	80	4	0.094

The two errors for reviewers not being cited were for citations to a book and a conference paper that did not have a DOI. All four errors in capturing self-citations were where the number captured was fewer than the true number, for example, we extracted 1 self-citation when the true number was 3.

References

[1] Yang, F.-X. et al. "RETRACTION: Origin of the distinct site occupations of H atom in hcp Ti and Zr/Hf". *International Journal of Hydrogen Energy* 91 (Nov. 2024), 933–941. DOI: 10.1016/j.ijhydene.2024.10.197.

[2] Seeber, M. et al. "Self-citations as strategic response to the use of metrics for career decisions". *Research Policy* 48.2 (2019), 478–491. DOI: https://doi.org/10.1016/j.respol.2017.12.004.

[3] Cranford, S. "C.R.E.A.M: Citations Rule Everything Around Me". *Matter* 2.6 (June 2020), 1343–1347. DOI: 10.1016/j.matt.2020.04.025.

[4] Burton, S. et al. "Cite me! Perspectives on coercive citation in reviewing". *Journal of Services Marketing* 38.7 (Sept. 2024), 809–815. DOI: 10.1108/jsm-08-2024-0387.

[5] Teixeira da Silva, J. A. "The ethics of peer and editorial requests for self-citation of their work and journal". *Medical Journal Armed Forces India* 73.2 (2017), 181–183. DOI: https://doi.org/10.1016/j.mjafi.2016.11.008.

[6] Committee on Publication Ethics. *Citation manipulation*. Jan. 2019. DOI: 10.24318/cope.2019.3.1. URL: http://dx.doi.org/10.24318/cope.2019.3.1.

[7] Wren, J. D., Valencia, A., and Kelso, J. "Reviewer-coerced citation: case report, update on journal policy and suggestions for future prevention". *Bioinformatics* 35.18 (Jan. 2019), 3217–3218. DOI: 10.1093/bioinformatics/btz071.

[8] Hamilton, D. G. et al. "Meta-Research: Journal policies and editors' opinions on peer review". *eLife* 9 (Nov. 2020), e62529. DOI: 10.7554/eLife.62529.

[9] Mehregan, M. and Moghiman, M. "The Unnoticed Issue of Coercive Citation Behavior for Authors". *Publishing Research Quarterly* 40.2 (June 2024), 164–168. DOI: 10.1007/s12109-024-09994-0.

[10] Martin, B. R. "Whither research integrity? Plagiarism, self-plagiarism and coercive citation in an age of research assessment". *Research Policy* 42.5 (June 2013), 1005–1014. DOI: 10.1016/j.respol.2013.03.011.

[11] Heneberg, P. "From Excessive Journal Self-Cites to Citation Stacking: Analysis of Journal Self-Citation Kinetics in Search for Journals, Which Boost Their Scientometric Indicators". *PLOS ONE* 11.4 (Apr. 2016), e0153730. DOI: 10.1371/journal.pone.0153730.

[12] Fong, E. A., Patnayakuni, R., and Wilhite, A. W. "Accommodating coercion: Authors, editors, and citations". *Research Policy* 52.5 (June 2023), 104754. DOI: 10.1016/j.respol.2023.104754.

[13] Singh Chawla, D. "Two-thirds of researchers report 'pressure to cite' in Nature poll". *Nature* (Oct. 2019). DOI: 10.1038/d41586-019-02922-9.

[14] Resnik, D. B., Gutierrez-Ford, C., and Peddada, S. "Perceptions of Ethical Problems with Scientific Journal Peer Review: An Exploratory Study". *Science and Engineering Ethics* 14.3 (Mar. 2008), 305–310. DOI: 10.1007/s11948-008-9059-4.

[15] Singh Chawla, D. "Elsevier investigates hundreds of peer reviewers for manipulating citations". *Nature* 573.7773 (Sept. 2019), 174. DOI: 10.1038/d41586-019-02639-9.

[16] Frey, B. S., Eichenberger, R., and Frey, R. L. "Editorial Ruminations: Publishing Kyklos". *Kyklos* 62.2 (Apr. 2009), 151–160. DOI: 10.1111/j.1467-6435.2009.00428.x.

[17] Schriger, D. L., Kadera, S. P., and von Elm, E. "Are Reviewers' Scores Influenced by Citations to Their Own Work? An Analysis of Submitted Manuscripts and Peer Reviewer Reports". *Annals of Emergency Medicine* 67.3 (2016), 401–406. DOI: https://doi.org/10.1016/j.annemergmed.2015.09.003.

[18] Stelmakh, I. et al. "Cite-seeing and reviewing: A study on citation bias in peer review". *PLOS ONE* 18.7 (July 2023), 1–16. DOI: 10.1371/journal.pone.0283980.

[19] Frandsen, T. F. and Nicolaisen, J. "Praise the bridge that carries you over: Testing the flattery citation hypothesis". *Journal of the American Society for Information Science and Technology* 62.5 (Mar. 2011), 807–818. DOI: 10.1002/asi.21503.

[20] Smith, R. "Peer Review: A Flawed Process at the Heart of Science and Journals". *Journal of the Royal Society of Medicine* 99.4 (2006), 178–182. DOI: 10.1177/014107680609900414.

[21] Bohorquez, N. G. et al. "Health and medical researchers are willing to trade their results for journal prestige: results from a discrete choice experiment". *Prometheus* (2025).

[22] Lee, C. J. et al. "Bias in peer review". *Journal of the American Society for Information Science and Technology* 64.1 (2013), 2–17. DOI: https://doi.org/10.1002/asi.22784.

[23] Schmidt, B. et al. "Ten considerations for open peer review". *F1000Research* 7 (June 2018), 969. DOI: 10.12688/f1000research.15334.1.

[24] Tennant, J. P. and Ross-Hellauer, T. "The limitations to our understanding of peer review". *Research Integrity and Peer Review* 5.1 (Apr. 2020). DOI: 10.1186/s41073-020-00092-1.

[25] Aczel, B. et al. "The present and future of peer review: Ideas, interventions, and evidence". *Proceedings of the National Academy of Sciences* 122.5 (Jan. 2025). DOI: 10.1073/pnas.2401232121.

[26] Vendé, B., Barberousse, A., and Ruphy, S. "From 2015 to 2023, eight years of empirical research on research integrity: a scoping review". *Research Integrity and Peer Review* 10.1 (Apr. 2025). DOI: 10.1186/s41073-025-00163-1.

[27] Currie, G. *Open Science: What is publish, review, curate?* 2024. URL: https://elifesciences.org/inside-elife/dc24a9cd/open-science-what-is-publish-review-curate.

[28] F1000Research. *Finding Article Reviewers*. 2025. URL: https://f1000research.com/for-authors/tips-for-finding-referees.

[29] F1000Research. *Guidelines For Article Reviewers*. 2025. URL: https://f1000research.com/for-referees/guidelines.

[30] Begg, C. B. and Berlin, J. A. "Publication Bias: A Problem in Interpreting Medical Data". *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 151.3 (1988), 419. DOI: 10.2307/2982993.

[31] Committee on Publication Ethics. *Ethical Guidelines for Peer Reviewers*. Sept. 2017. DOI: 10.24318/cope.2019.1.9.

[32] Culbert, J. H. et al. "Reference coverage analysis of OpenAlex compared to Web of Science and Scopus". *Scientometrics* 130.4 (Apr. 2025), 2475–2492. DOI: 10.1007/s11192-025-05293-3.

[33] Priem, J., Piwowar, H., and Orr, R. *OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts*. 2022. arXiv: 2205.01833 [cs.DL].

[34] Massimo, A. et al. "openalexR: An R-Tool for Collecting Bibliometric Data from OpenAlex". *The R Journal* 15 (4 2024), 167–180. DOI: 10.32614/RJ-2023-089.

[35] Burnham, K. and Anderson, D. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer New York, 2002. ISBN: 9780387953649.

[36] Bland, J. M. and Altman, D. G. "Statistics notes: Matching". *BMJ* 309.6962 (Oct. 1994), 1128. DOI: 10.1136/bmj.309.6962.1128.

[37] Hosmer, D., Lemeshow, S., and Sturdivant, R. *Applied Logistic Regression*. Wiley Series in Probability and Statistics. Wiley, 2013. ISBN: 9780470582473.

[38] Zou, H. and Hastie, T. "Regularization and Variable Selection Via the Elastic Net". *Journal of the Royal Statistical Society Series B: Statistical Methodology* 67.2 (Mar. 2005), 301–320. DOI: 10.1111/j.1467-9868.2005.00503.x.

[39] Tay, J. K., Narasimhan, B., and Hastie, T. "Elastic Net Regularization Paths for All Generalized Linear Models". *Journal of Statistical Software* 106.1 (2023), 1–31. DOI: 10.18637/jss.v106.i01.

[40] Barnett, A. *F1000Research – reviewer citation study*. May 2024. URL: https://aspredicted.org/rn8vg.pdf.

[41] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2024. URL: https://www.R-project.org/.

[42] Barnett, A. *Code and data for the analysis of the association between citations and peer review recommendations*. 2025. DOI: 10.5281/zenodo.16551814. URL: https://github.com/agbarnett/cited_reviewers.

[43] *S D. Peer Review Report For: COVID-19 Vaccine: Predicting Vaccine Types and Assessing Mortality Risk Through Ensemble Learning Algorithms [version 2; peer review: 2 approved, 2 approved with reservations]*. 2024. DOI: 10.5256/f1000research.153740.r257039.

[44] Oviedo-García, M. Á. "The review mills, not just (self-)plagiarism in review reports, but a step further". *Scientometrics* 129.9 (Aug. 2024), 5805–5813. DOI: 10.1007/s11192-024-05125-w.

[45] Anderson, M. S. et al. "The Perverse Effects of Competition on Scientists' Work and Relationships". *Science and Engineering Ethics* 13.4 (Nov. 2007), 437–461. DOI: 10.1007/s11948-007-9042-5.

[46] Seeber, M. "How do journals of different rank instruct peer reviewers? Reviewer guidelines in the field of management". *Scientometrics* 122.3 (Jan. 2020), 1387–1405. DOI: 10.1007/s11192-019-03343-1.

[47] Thombs, B. D. and Razykov, I. "A solution to inappropriate self-citation via peer review". *Canadian Medical Association Journal* 184.16 (Oct. 2012), 1864. DOI: 10.1503/cmaj.120597.

[48] Ross-Hellauer, T., Deppe, A., and Schmidt, B. "Survey on open peer review: Attitudes and experience amongst editors, authors and reviewers". *PLOS ONE* 12.12 (Dec. 2017), 1–28. DOI: 10.1371/journal.pone.0189311.

[49] Taylor and Francis. *Anonymous peer review: How to make your article ready for double-anonymous peer review*. 2025. URL: https://authorservices.taylorandfrancis.com/publishing-your-research/peer-review/anonymous-peer-review/.

[50] Bergstrom, C. T. and Bak-Coleman, J. "AI, peer review and the human activity of science". *Nature* (June 2025). DOI: 10.1038/d41586-025-01839-w.

[51] Thombs, B. D. et al. "Potentially coercive self-citation by peer reviewers: A cross-sectional study". *Journal of Psychosomatic Research* 78.1 (2015), 1–6. DOI: https://doi.org/10.1016/j.jpsychores.2014.09.015.

[52] Peebles, E., Scandlyn, M., and Hesp, B. R. "A retrospective study investigating requests for self-citation during open peer review in a general medicine journal". *PLOS ONE* 15.8 (Aug. 2020), 1–9. DOI: 10.1371/journal.pone.0237804.

[53] Sugimoto, C. R. and Cronin, B. "Citation gamesmanship: testing for evidence of ego bias in peer review". *Scientometrics* 95.3 (Sept. 2012), 851–862. DOI: 10.1007/s11192-012-0845-z.

[54] Fong, E. A. and Wilhite, A. W. "Authorship and citation manipulation in academic research". *PLOS ONE* 12.12 (Dec. 2017), 1–34. DOI: 10.1371/journal.pone.0187394.

[55] Wilhite, A. W. and Fong, E. A. "Coercive Citation in Academic Publishing". *Science* 335.6068 (Feb. 2012), 542–543. DOI: 10.1126/science.1212540.

[56] Ho, R. C.-M. et al. "Views on the peer review system of biomedical journals: an online survey of academics from high-ranking universities". *BMC Medical Research Methodology* 13.1 (June 2013). DOI: 10.1186/1471-2288-13-74.

[57] Pavlovic, V. et al. "How accurate are citations of frequently cited papers in biomedical literature?" *Clinical Science* 135.5 (Mar. 2021), 671–681. DOI: 10.1042/cs20201573.

[58] Clarke, R. et al. "Underestimation of Risk Associations Due to Regression Dilution in Long-term Follow-up of Prospective Studies". *American Journal of Epidemiology* 150.4 (Aug. 1999), 341–353. DOI: 10.1093/oxfordjournals.aje.a010013.

[59] Fire, M. and Guestrin, C. "Over-optimization of academic publishing metrics: observing Goodhart's Law in action". *GigaScience* 8.6 (May 2019). ISSN: 2047-217X. DOI: 10.1093/gigascience/giz053. URL: http://dx.doi.org/10.1093/gigascience/giz053.

[60] Tahamtan, I. and Bornmann, L. "What do citation counts measure? An updated review of studies on citations in scientific documents published between 2006 and 2018". *Scientometrics* 121.3 (Sept. 2019), 1635–1684. DOI: 10.1007/s11192-019-03243-4.

[61] Rooyen, S. van et al. "Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial". *BMJ* 318.7175 (Jan. 1999), 23–27. DOI: 10.1136/bmj.318.7175.23.

[62] Gomez, C. J., Herman, A. C., and Parigi, P. "Leading countries in global science increasingly receive more citations than other countries doing similar research". *Nature Human Behaviour* 6.7 (May 2022), 919–929. DOI: 10.1038/s41562-022-01351-5.

[63] Campos-Arceiz, A., Primack, R. B., and Koh, L. P. "Reviewer recommendations and editors' decisions for a conservation journal: Is it just a crapshoot? And do Chinese authors get a fair shot?" *Biological Conservation* 186 (2015), 22–27. DOI: https://doi.org/10.1016/j.biocon.2015.02.025.

[64] Royston, P., Ambler, G., and Sauerbrei, W. "The use of fractional polynomials to model continuous risk variables in epidemiology". *International Journal of Epidemiology* 28.5 (Oct. 1999), 964–974. DOI: 10.1093/ije/28.5.964.

[65] Therneau, T. and Grambsch, P. *Modeling Survival Data: Extending the Cox Model*. Statistics for Biology and Health. Springer New York, 2013. ISBN: 9781475732948.