Gender and international diversity improves equity in peer review

Dakota Murray[1], Kyle Siler[2], Vincent Larivière[3], Wei Mun Chan[4], Andrew M. Collings[4], Jennifer Raymond[5◑], Cassidy R. Sugimoto[1◑*]

1 School of Informatics Computing, and Engineering, Indiana University Bloomington, Bloomington, Indiana, U.S.A.
2 Copernicus Institute for Sustainable Development, Utrecht University, Utrecht, The Netherlands
3 École de bibliothéconomie et des sciences de l'information, Université de Montréal, Montréal, Québec, Canada
4 eLife Sciences Publishing Ltd., Cambridge, United Kingdom
5 Department of Neurobiology, Stanford University, Stanford, California, U.S.A.

◑These authors contributed equally to this work.
* corresponding author: sugimoto@indiana.edu

Abstract

The robustness of scholarly peer review has been challenged by evidence of disparities in publication outcomes based on author's gender and nationality. To address this, we examine the peer review outcomes of 23,873 initial submissions and 7,192 full submissions that were submitted to the biosciences journal *eLife* between 2012 and 2017. Women and authors from nations outside of North America and Europe were underrepresented both as gatekeepers (editors and peer reviewers) and last authors. We found a homophilic interaction between the demographics of the gatekeepers and authors in determining the outcome of peer review; that is, gatekeepers favor manuscripts from authors of the same gender and from the same country. The acceptance rate for manuscripts with male last authors was significantly higher than for female last authors, and this gender inequity was greatest when the team of reviewers was all male; mixed-gender gatekeeper teams lead to more equitable peer review outcomes. Similarly, manuscripts were more likely to be accepted when reviewed by at least one gatekeeper with the same national affiliation as the corresponding author. Our results indicated that homogeneity between author and gatekeeper gender and nationality is associated with the outcomes of scientific peer review. We conclude with a discussion of mechanisms that could contribute to this effect, directions for future research, and policy implications. Code and anonymized data have been made available at https://github.com/murrayds/elife-analysis

Author summary

Peer review, the primary method by which scientific work is evaluated and developed, is ideally a fair and equitable process, in which scientific work is judged solely on its own merit. However, the integrity of peer review has been called into question based on evidence that outcomes often differ between between male and female authors, and between authors in different countries. We investigated such a disparity at the biosciences journal *eLife*, by analyzing the author and gatekeepers (editors and peer reviewers) demographics and review outcomes of all submissions between 2012 and 2017. We found evidence of disparity in outcomes that disfavored women and those outside of North America and Europe, and that these groups were underrepresented among authors and gatekeepers. The gender disparity was greatest when reviewers were all male; mixed-gender reviewer teams lead to more equitable outcomes. Similarly, manuscripts were more likely to be accepted when reviewed by at least one gatekeeper from the same country as the corresponding author. Our results indicated that gatekeeper characteristics are associated with the outcomes of scientific peer review. We discuss mechanisms that could contribute to this effect, directions for future research, and policy implications.

Introduction

Peer review is foundational to the development, gatekeeping, and dissemination of research, while also underpinning the professional hierarchies of academia. Normatively, peer review is expected to follow the ideal of "universalism" [1], where scholarship is judged solely on its intellectual merit. However, confidence in the extent to which peer review accomplishes the goal of promoting the best scholarship has been eroded by questions about whether social biases [2], based on or correlated with the characteristics of the scholar, could also influence outcomes of peer review [3, 4]. This challenge to the integrity of peer review has prompted an increasing number of funding agencies and journals to assess the disparities and potential influence of bias in their peer review processes.

Several terms are often conflated in the discussion on bias in peer review. We use the term *disparities* to refer to unequal composition between groups, *inequities* to characterize unequal outcomes, and *bias* to refer to the degree of impartiality in judgment. Disparities and inequities have been widely studied in scientific publishing, most notably in regards to gender and country of affiliation. Globally, women account for about 30 percent of scientific authorship [5] and are underrepresented in the scientific workforce [6, 7]. Articles authored by women are disproportionately underrepresented in the most prestigious and high-profile scientific journals [8–13]. Moreover, developed countries dominate the production of highly-cited publications [14, 15].

The underrepresentation of authors from certain groups may reflect differences in submission rates, or it may reflect differences in success rates during peer review (percent of submissions accepted). Analyses of success rates have yielded mixed results in terms of the presence and magnitude of such inequities. Some analyses have found lower success rates for female-authored papers [16, 17] and grant applications [18, 19], while other studies have found no gender differences in review outcomes (for examples, see [20–23]). Inequities in journal success rates based on authors' nationalities have also been documented, with reports that authors from English-speaking and scientifically-advanced countries have higher success rates [24, 25]; however, other studies found no evidence that the language or country of affiliation of an author influences peer review outcomes [25–28]. These inconsistencies could be explained by several factors, such as the contextual factors of the studies and the variations in research design and sample size.

The nature of bias and its contribution to inequities in scientific publishing is highly controversial. Implicit bias—the macro-level social and cultural stereotypes that can subtly influence everyday interpersonal judgements and thereby produce and perpetuate status inequalities and hierarchies [29, 30]—has been suggested as a possible mechanism to explain differences in peer review outcomes based on socio-demographic and professional characteristics [3, 31]. When faced with uncertainty—which is quite common in academia—people often weight the social status and other ascriptive characteristics of others to help make decisions [32]. Hence, scholars are more likely to consider particularistic characteristics (e.g., gender, institutional status) of an author under conditions of uncertainty [33, 34], such as at the frontier of new scientific knowledge [35]. However, given the stratification of scholars within institutions and across countries, it can be difficult to pinpoint the nature of a potential bias. For example, women are underrepresented in prestigious educational institutions [36–38], which conflates gender and prestige biases. These institutional differences can be compounded by gendered differences in age, professional seniority, research topic, and access to top mentors [39]. Another potential source of bias is what [40] dubbed *cognitive particularism*, where scholars harbor preferences for work and ideas similar to their own [41]. Evidence of this process has been reported in peer review at one journal, in the reciprocity (i.e., correspondences between patterns of recommendations received by authors and patterns of recommendations given by reviewers in the same social group) between authors and reviewers of the same race and gender [42] (see also [43, 44]). Reciprocity can exacerbate or mitigate existing inequalities in science. If the work and ideas favored by gatekeepers are unevenly distributed across author demographics, this could be conducive to Matthew Effects [1], whereby scholars accrue accumulative advantages via *a priori* status privileges. Consistent with this, inclusion of more female reviewers was reported to attenuate biases that disfavor women in the awarding of Health RO1 grants at the National Institute of Health [17]. However, an inverse reaction was found by [45] in the evaluation of candidates for professorships: when female evaluators were present, male evaluators became less favorable toward female candidates. Thus the nature and potential impact of cognitive biases during peer review are multiple and complex.

Another challenge is to disentangle the contribution of bias during peer review from factors external to the review process that could influence success rates. For example, there are gendered differences in access to funding, domestic responsibilities, and cultural expectations of career preferences and ability [46, 47] that may adversely impact manuscript preparation and submission. Furthermore, women have been found to be less likely to compete [48] and hold themselves to higher standards [49], hence they may self-select a higher quality of work for submission to prestigious journals. At the country level, disparities in peer review outcomes could reflect structural factors related to a nation's scientific investment [14, 50], publication incentives [51, 52], local challenges [53], and research culture [54], all of which could influence the actual and perceived quality of submissions from different nations. Because multiple factors external to the peer review process can influence peer review outcomes, unequal success rates for authors with particular characteristics do not necessarily reflect bias, conversely, equal success rates do not necessarily reflect a lack of bias.

Several approaches have been applied in an attempt to identify and analyze bias. One quasi-experimental approach was to compare outcomes in anonymized (double-blind) and non-anonymized (single-blind) peer review. The results of these studies demonstrates that double-blind review yields more equitable outcomes [31] and mitigates inequities that favor famous authors, elite institutions [55–57], and those from high-income and English-speaking nations [26]. However, although double-blind review is generally viewed positively by the scientific community [58, 59], this process has not been widely used in peer review in the biosciences, and authors rarely opted-in when offered [60]. Therefore, the experimental approach has not resolved debate over the role of bias in the outcomes of peer review in scholarly publishing.

Here, we use an alternative approach to assess the extent to which gender and national disparities manifest in peer review outcomes at *eLife*—an open-access journal in the life and biomedical sciences. In particular, we study the extent to which the magnitude of these disparities vary across different gender and national compositions of gatekeeper teams. Peer review at *eLife* differs from other traditional forms of peer review used in the life sciences in that review at *eLife* is done through deliberation between reviewers (usually three in total) on an online platform. Previous studies have shown that deliberative scientific evaluation are influenced by social dynamics between evaluators [61, 62]. Therefore, we assessed the extent to which the composition of the reviewer teams relates to peer review outcomes. Using all research papers (Research Articles, Short Reports, and Tools and Resources) submitted between 2012 and 2017 (n=23,879), we investigated the extent to which an interaction manifests between the gender and nationality of authors (first, last, and corresponding) and gatekeepers (editors and invited peer reviewers), similar to the approach used by [2]. We acknowledge that inequity in success rates could result from a variety of factors unrelated to the peer review process. However, if the outcomes vary significantly based on the demographic characteristics of the reviewers in relation to authors, we contend that this provides evidence of potential bias in the peer review process.

Data and methods

Consultative peer review and *eLife*

Founded in 2012 by the Howard Hughes Medical Institute (United States), the Max Planck Society (Germany), and the Wellcome Trust (United Kingdom), *eLife* is an open-access journal that publishes research in the life and biomedical sciences. Manuscripts submitted to *eLife* progress through several stages. In the first stage, the manuscript is assessed by a Senior Editor, who may confer with one or more Reviewing Editors and decide whether to reject the manuscript or encourage the authors to provide a full submission. In May 2018, *eLife* had 45 Senior Editors, including the Editor-in-Chief and three Deputy Editors, and 339 Reviewing Editors, all of whom were active scientists. When a full manuscript is submitted, the Reviewing Editor recruits a small number of peer reviewers (typically two or three) to write reports on the manuscript. The Reviewing Editor is encouraged to serve as one of the peer reviewers; in our sample, the Reviewing Editor was listed as a peer reviewer for 58.9 percent of full submissions. When all individual reports have been submitted, both the Reviewing Editor and peer reviewers discuss the manuscript and their reports using a private online chat system hosted by *eLife*. At this stage the identities of the Reviewing Editor and peer reviewers are known to one another. If the consensus of this group is to reject the manuscript, all the reports are usually sent to the authors. If the consensus is that the manuscript requires revision, the Reviewing Editor and additional peer reviewers agree on the essential points that need to be addressed before the paper can be accepted. In this case, a decision letter outlining these points is sent to the authors (and the original reports by the peer reviewers are not usually sent in their entirety to the authors). When a manuscript is accepted, the decision letter and the authors' response are published along with the manuscript. The name of the Reviewing Editor is also published. Peer reviewers can also choose to have their name published. This process has been referred to as consultative peer review (see [63, 64] for a more in-depth description of the *eLife* peer-review process). Consultative peer review provides a unique context for analyzing deliberation and social valuation in professional groups.

Data

We retrieved metadata for research papers submitted to *eLife* between its inception in 2012 and mid-September, 2017 (n=23,873). Submissions fell into three main categories: 20,945 Research Articles (87.7 percent), 2,186 Short Reports (9.2 percent), and 742

Tools and Resources (3.1 percent). Not included in this total were six Scientific Correspondence articles, which were excluded because they follow a distinct review process. Each record potentially included four submissions—an initial submission, full submission, and up to two revision submissions (though in some cases manuscripts remain in revision even after two revised submissions). Fig 1 depicts the flow of all 23,873 manuscripts through each review stage. The majority, 70.0 percent, of initial submissions for which a decision was made were rejected. Only 7,111 manuscripts were encouraged for a full submission. A total of 7,192 manuscripts were submitted as full submission; the number is slightly larger than encouraged initial submissions due to appeals of initial decisions and other special circumstances. Most full submissions, 52.4 percent (n = 3,767), received a decision of revise, while another 43.9 percent (n = 3,154) were rejected. A small number of full submissions (n = 54) were accepted without any revisions. On average, full submissions that were ultimately accepted underwent 1.23 revisions and, within our dataset, 3,426 full submissions were eventually accepted to be published. A breakdown of the number of revisions requested before a final decision is made, by gender and nationality of the last author, is provided in S1 Fig. On the date on which data was collected (mid-September, 2017), a portion of initial submission (n = 147) and full submissions (n = 602) remained in various stages of processing and deliberation (without final decisions). Another portion of initial and full submissions (n = 619) appealed their decision, causing some movement from decisions of "Reject" to decisions of "Accept" or "Revise".

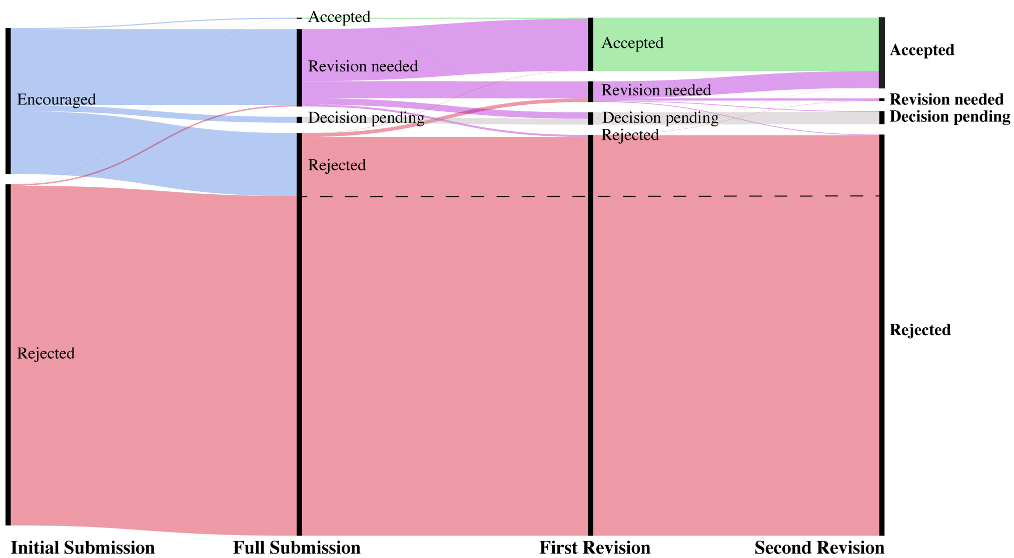


Fig 1. Flow of papers through the *eLife* review process.
Begins with an initial submission and initial decision of encourage or reject, and then leading to the first full review and subsequent rounds of revision. "Encouraged", "Accepted", "Rejected" and "Revision needed" represent the decisions made by *eLife* editors and reviewers at each submission stage. A portion of manuscripts remained in various stages of processing at the time of data collection—these manuscripts were labeled as "Decision pending". The status of manuscripts after the second revision is the final status that we consider in the present data. The dashed line delineates full submissions from rejected initial submissions.

The review process at *eLife* was highly selective, and became more selective over time. Fig 2 shows that while the total count of manuscripts submitted to *eLife* has rapidly increased since the journal's inception, the count of encouraged initial submissions and accepted full submissions has grown more slowly. The encourage rate (percentage of initial submissions encouraged to submit full manuscripts) was 44.6 percent in 2012, and dropped to 26.6 percent in 2016. The overall accept rate (percentage of initial submissions eventually accepted) began at 27.0 percent in 2012 and decreased to 14.0 percent in 2016. The accept rate (the percentage of accepted full submissions) was 62.4 percent in 2012 and decreased to 53.0 percent in 2016. While only garnering 307 submissions in 2012, *eLife* accrued 8,061 submissions in 2016. In the present analysis we considered the outcomes of all manuscripts without respect to submission year, though we note that data was skewed to the large portion of manuscripts published most recently.

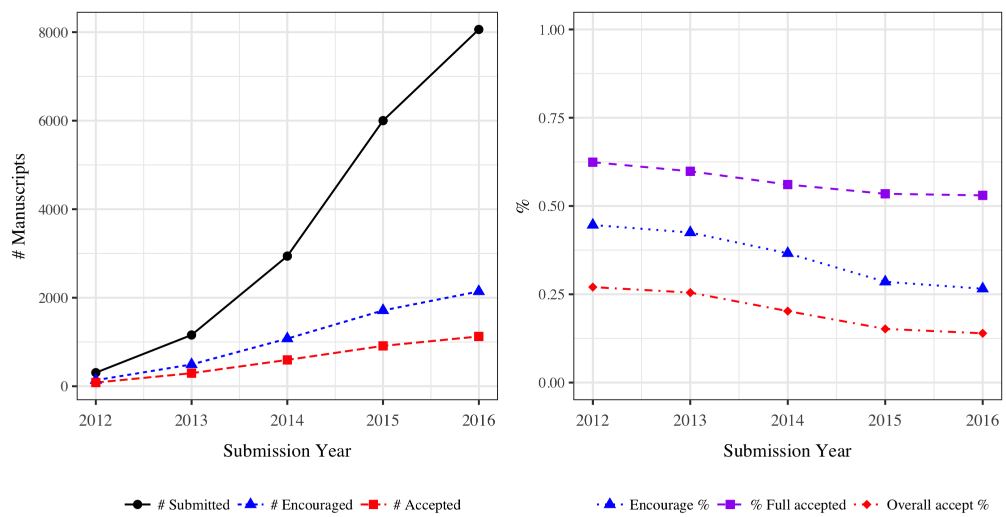


Fig 2. Submissions and selectivity of *eLife* over time.
Left: Yearly count of initial submissions, encouraged initial submissions, and accepted full submissions to *eLife* between 2012 and 2016; Right: rate of initial submissions encouraged (Encourage %), rate of full submissions accepted (% Full accepted) and rate of initial submissions accepted (Overall accept %) between 2012 and 2016. Submissions during the year of 2017 were excluded because we do not have sufficient data for full life-cycle of these manuscripts.

In addition to authorship data, we obtained information about the gatekeepers involved in the processing of each submission. In our study, we define gatekeepers to include any Senior Editor or Reviewing Editor at *eLife* or invited peer reviewer involved in the review of at least one initial or full submission between 2012 and mid-September 2017. Gatekeepers at *eLife* often serve in multiple roles, for example, acting as both a Reviewing Editor and peer reviewer on a given manuscript. For initial submissions, we had data on the corresponding author of the manuscript and the Senior Editor tasked with making the decision. For full submissions we had data on the corresponding author, first author, last author, Senior Editor, Reviewing Editor, and members of the team of peer reviewers. Data for each individual included their stated name, institutional affiliation, and country of affiliation. A small number of submissions were removed, such as cases where a paper had a first but no last author or papers which did not have a valid submission type. Country names were manually disambiguated (for example, normalized names such as "USA" to "United States" and "Viet Nam" to "Vietnam").

Full submissions included 6,669 distinct gatekeepers, 6,694 distinct corresponded authors, 6,691 distinct first authors, and 5,580 distinct last authors. Authors were also likely to appear on multiple manuscripts and may hold a different authorship role in each: in 26.5 percent of full submissions the corresponding author was also the first author, while in 71.2 percent of submissions the corresponding author was also the last author. We did not have access to the full authorship list that included middle authors. Note that in the biosciences, the last author is typically the most senior researcher involved [65] and responsible for more conceptual work, whereas the first author is typically less senior and performs more of the scientific labor (such as lab work, analysis, etc.) to produce the study [66–68].

Gender assignment

Gender variables for authors and gatekeepers were coded using an updated version of the algorithm developed in [5]. This algorithm used a combination of the first name and country of affiliation to assign each author's gender on the basis of several universal and country-specific name-gender lists (e.g., United States Census). This list of names was complemented with an algorithm that searches Wikipedia for pronouns associated with names. This new list was validated by applying it to a dataset of names with known gender. We used data collected from RateMyProfessor.com, a website containing anonymous student-submitted ratings and comments for professors, lecturers, and teachers for professors at United States, United Kingdom and Canadian universities. We limited the dataset to only individuals with at least five comments, and counted the total number of gendered pronouns that appear in comments; if the total of one gendered-pronoun type was at least the square of the other, then we assigned the gender of the majority pronoun to the individual. To compare with pronoun-based assignment, we assigned gender using the previously detailed first-name based algorithm. In total, there were 384,127 profiles on RateMyProfessor.com that had at least five comments and for whom pronouns indicated a gender. Our first name-based algorithm assigned a gender of male or female to 91.26 percent of these profiles. The raw match-rate between these two assignments was 88.6 percent. Of those that were assigned a gender, our first name-based assignment matched the pronoun assignment in 97.1 percent of cases, and 90.3 percent of distinct first names. While RateMyProfessor.com and the authors submitting to *eLife* represent different populations (RateMyProfessor.com being biased towards teachers in the United States, United Kingdom, and Canada), the results of this validation provide some credibility to the first-name based gender assignment used here. We also manually identified gender for three Senior Editors and 24 Reviewing Editors for whom our algorithm did not assign gender by searching for them on the web and inspected resulting photos in order to determine if the individual was presenting as male or female.

Through the combination of manual efforts and our first-name based gender-assignment algorithm, we assigned a gender of male or female to 92.3 percent (n = 34,333) of the 37,195 name/role combinations that appear in our dataset. 26.0 percent (n = 9,675) were assigned a gender of female, 66.3 percent (n = 24,658) were assigned a gender of male, while the remaining 7.7 percent (n = 2,862) were assigned no gender. This gender distribution roughly matches the gender distribution observed globally across scientific publications [5].

Analysis

When comparing peer review outcomes between groups, we used χ^2 tests of independence. We maintain the convention of 0.05 as the threshold of statistical significance, though we also report significance levels less than or equal to 0.1 as marginally significant. When visualizing our results, we superimposed the 95$^{\text{th}}$ percentile sample proportion confidence intervals. When comparing groups based on gender, we excluded submissions for which no gender could be identified. Data processing, statistical testing, and visualization was performed using R version 3.4.2 and RStudio version 1.1.383.

Results

Gatekeeper representation

We first analyzed whether the gender and national affiliations of the population of gatekeepers at *eLife* was similar to that of the authors of initial and full submissions (Fig 2). The population of gatekeepers was primarily comprised of invited peer reviewers, as there were far fewer Senior and Reviewing Editors. A gender breakdown by type of gatekeeper has been provided in S1 Table, and a national breakdown is provided in S2 Table.

Fig 3 illustrates the gender and national demographics of authors and gatekeepers at *eLife*. The population of gatekeepers at *eLife* was largely male. Only 20.6 percent (n = 1,372) of gatekeepers were identified as female, compared with 26.4 percent (n = 4,803) of corresponding authors (includes authors of initial submissions), 33.6 percent (n = 2,256) of first authors, and 22.2 percent (n = 1,243) of last authors. The difference between the gender composition of gatekeepers and authors was statistically significant for corresponding authorship, $\chi^2(1, n = 16,774) = 465.9, p < 0.0001$; first authorship, $\chi^2(1, n = 6,087) = 837.6, p < 0.0001$; and last authorship, $\chi^2(1, n = 5,162) = 16.4$, $p < 0.0001$. Thus, the gender proportions of gatekeepers at *eLife* was male-skewed in comparison to the authorship profile.

The population of gatekeepers at *eLife* was heavily dominated by those from North America, who constitute 59.9 percent (n = 3,992). Gatekeepers from Europe were the next most represented, constituting 32.4 percent (n = 2,161), followed by Asia with 5.7 percent (n = 379). Individuals from South America, Africa, and Oceania each made up less than two percent of the population of gatekeepers. As with gender, we identified significant differences between the international composition of gatekeepers and that of the authors. Gatekeepers from North America were over-represented whereas gatekeepers from Asia and Europe were under-represented compared to the population of corresponding authors, $\chi^2(5, n = 18,191) = 6904.6, p < 0.0001$, first authors, $\chi^2(5, n = 6,670) = 480.4, p < 0.0001$, and last authors, $\chi^2(5, n = 5,564) = 428.2$, $p < 0.0001$. The international representation of gatekeepers was most similar to first and last authorship, and least similar to corresponding authorship. This likely resulted from the fact that our population of corresponding authors included initial submissions, which tend to be more internationally diverse than full submissions, for which we had information about first and last authors as well as corresponding authors.

Authorship, Gender, and Outcomes

Male authorship dominated *eLife* submissions: men accounted for 76.8 percent (n = 5,113) of gender-identified last authorships and 74.0 percent (n = 4,913) of gender-identified corresponding authorships of full submissions (see S2 Fig). First authorship of full submissions was closest to gender parity, although still skewed towards male authorship at 63.2 percent (n = 4,125).

We found small but statistically significant gender inequity favoring men in the outcomes of each stage of the review processes. The percentage of initial submissions encouraged was higher for male corresponding authors—30.6 to 28.6 percent, $\chi^2(1, n = 21,841) = 7.79, p < 0.01$ (see S2 Fig). Likewise, the percentage of full submissions accepted was higher for male corresponding authors—53.4 to 50.4 percent $\chi^2(1, n = 6,013) = 4.0, p < 0.05$. The gender disparity at each stage of the review

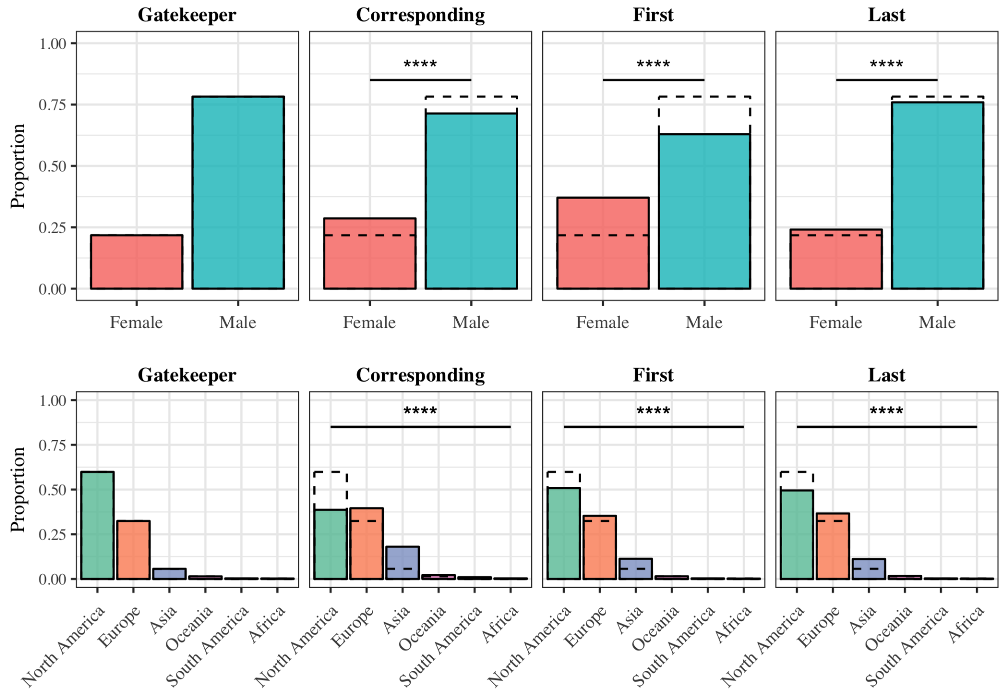


Fig 3. Gender and nationality demographics of authors and gatekeepers at *eLife*.
Top: proportion of identified men and women in the populations of distinct gatekeepers (Senior Editors, Reviewing Editors, and peer reviewers) and of the populations of distinct corresponding authors, first authors, and last authors; percentages exclude those for whom gender could be identified. Bottom: proportion of people with national affiliations within each of six continents in the population of distinct gatekeepers, and for the population of distinct corresponding, first, and last authors. Black dashed lines overlaid on authorship graphs indicate the proportion of gatekeepers within that gendered or continental category. Asterisks indicate the significance level of χ^2 tests of independence comparing the frequency of gender or continents between gatekeepers and each authorship type. "****" = $p < 0.0001$; "ns" = $p > 0.1$.

process yielded significantly higher overall accept rates (the percentage of initial submissions eventually accepted) for male corresponding authors (15.4 percent) compared with female corresponding authors (13.6 percent), $\chi^2(1, n = 21,217) = 10.5, p < 0.01$ (see S2 Fig).

Fig 4 shows the gendered acceptance rates of full submissions for corresponding, first, and last author. The greatest gender disparity in success rates was observed for last authors. The accept rate of full submissions was 3.7 percentage points higher for male last authors—53.4 to 49.7 percent, $\chi^2(1, n = 6,035) = 5.70, p < 0.05$. There was no significant relationship between the gender of the first author and percentage of full submissions accepted, $\chi^2(1, n = 5,913) = 0.42, p > 0.1$. Differences may be present at the intersection of gender and national affiliation (see S3 Fig), though the data was not sufficient to support statistically valid conclusions.

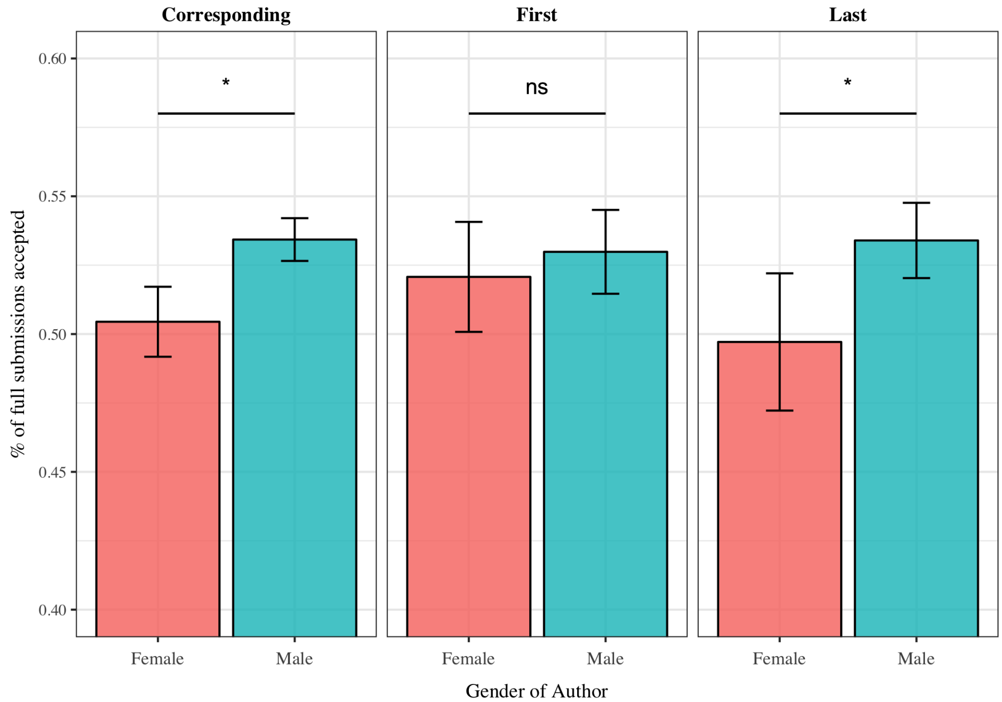


Fig 4. Peer review outcomes by gender of last author
Percentage of full submissions that were accepted, shown by the gender of the corresponding author, first author, and last author. Authors whose gender was unknown were excluded from analysis. See S2 Fig for an extension of this figure including submission rates, encourage rates, and overall acceptance races. Vertical error bars indicate 95$^{\text{th}}$ confidence intervals of the proportion of submitted, encouraged, and accepted initial and full submissions. Asterisks indicate significance level of χ^2 tests of independence of frequency of encourage and acceptance by gender; "*" $= p < 0.05$; "ns" $= p > 0.1$.

Gender homogeneity and peer review outcome

To examine the relationship between author-gatekeeper gender homogeneity on review outcomes, we analyzed the gender composition of the gatekeepers and authors of full submissions. Each manuscript was assigned a reviewer composition category of *all-male*, *all-female*, *mixed*, or *uncertain*. Reviewer teams labeled *all-male* and *all-female* were teams for which we could identify a gender for every member, and for which all genders were identified as either male or female, respectively. Teams labeled as *mixed* were those teams where we could identify a gender for at least two members, and which had at least one male and at least one female peer reviewer. Teams labeled as *uncertain* were those teams for which we could not assign a gender to every member and which were not mixed. A full submission is typically reviewed by two to three peer reviewers, which may or may not include the Reviewing Editor. However, the Reviewing Editor is likely to some degree always involved in the review process of a manuscript, and so we always considered the Reviewing Editor as a member of the reviewing team. Of 7,912 full submissions, a final decision of accept or reject was given for 6,590 during the dates analyzed; of these, 40.9 percent (n = 2,696) were reviewed by all-male teams, 1.2 percent (n = 81) by all-female teams, and 49.0 percent (n = 3,226) by mixed-gender teams; the remaining 587 reviewer teams were classified as uncertain.

Fig 4 illustrates higher acceptance rates for full submissions from male corresponding and last authors (submissions with authors of unidentified gender excluded). Fig 5 shows that this disparity manifested largely from instances when the reviewer team was all male. When all reviewers were male, the acceptance rate of full submissions was significantly higher for male compared to female last authors (Fig 5; $\chi^2(1, n = 2,472) = 6.6, p < 0.05$) and corresponding authors (S3 Fig; $\chi^2(1, n = 2,472) = 4.5, p < 0.05$). For mixed-gender reviewer teams, the disparity in author success rates by gender was smaller and non-significant. All-female reviewer teams were rare (only 81 of 6,509 processed full submissions). In the few cases of all-female reviewer teams, there was a higher acceptance rate for female last, corresponding, and first authors; however, this difference was not statistically significant, and the number of observations was too small to draw conclusions. There was no significant relationship between first authorship gender and acceptance rates, regardless of the gender composition of the reviewer team (see S3 Fig). In summary, we found that full submissions with male corresponding and last authors were more often accepted when they were reviewed by a team of gatekeepers consisting only of men; greater parity in outcomes was observed when gatekeeper teams contained both men and women. We refer to this favoring by reviewers of authors sharing the same gender as homophily.

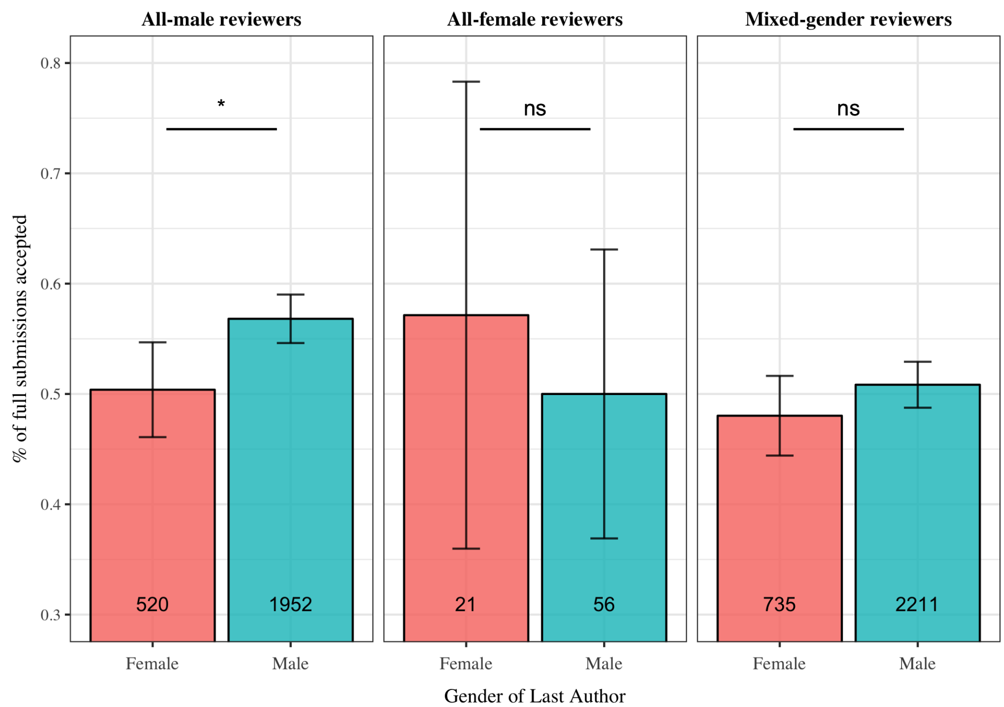


Fig 5. Interaction between author gender and gender-composition of gatekeepers.
Percentage of full submissions that were accepted, shown by the gender of the last author, and divided by the gender composition of the peer reviewers. Text at the base of each bar indicate the number full submissions within each category of reviewer team and authorship gender. Vertical error bars indicate 95$^{\text{th}}$ percentile confidence intervals of the proportion of accepted full submissions. See S3 Fig for all combinations of authorship and gatekeeper composition. Asterisks indicate significance level of χ^2 tests of independence on frequency of acceptance by gender of author given each team composition; "ns" indicates no observed statistical significance. "*" = $p < 0.05$; "ns" = $p > 0.1$.

Country of affiliation and peer review outcome

Fig 6 shows the proportions and rates of manuscripts submitted, encouraged, and accepted to *eLife* for corresponding authors originating from the eight most prolific countries (in terms of initial submissions). Manuscripts with corresponding authors from these eight countries accounted for a total of 73.9 percent of all initial submissions, 81.2 percent of all full submissions, and 86.5 percent of all accepted publications. Many countries were underrepresented in full and accepted submissions compared to their submissions. For example, while papers with Chinese corresponding authors accounted for 6.9 percent of initial submissions, they comprised only 3.0 percent of full and 2.4 percent of accepted submissions. The only countries that were overrepresented—making up a greater portion of full and accepted submissions than expected given their initial submissions—were the United States, United Kingdom, and Germany. In particular, corresponding authors from the United States made up 35.8 percent of initial submissions, yet constituted 48.5 percent of full submissions and the majority (54.9 percent) of accepted submissions.

Each stage of review contributed to the disparity of national representation between initial, full, and accepted submissions, with manuscripts from the United States, United Kingdom, and Germany more often encouraged as initial submissions, and accepted as full submissions. Figure 6 shows that initial submissions with a corresponding author from the United States were the most likely to be encouraged (39.2 percent), followed by the United Kingdom (31.7 percent) and Germany (29.3 percent). By contrast, manuscripts with corresponding authors from Japan, Spain, and China were least likely of these eight to be encouraged (21.4, 16.7, and 12.6 percent, respectively). These differences narrowed somewhat for full submissions; the accept rate for full submissions with corresponding authors from the U.S. was the highest (57.6 percent), though more similar to the United Kingdom and France than encourage rates. Full submissions from China, Spain, and Japan had the lowest acceptance rates of these eight countries

Country homogeneity and peer review outcomes

We also investigated the relationship between peer review outcomes and the presence of nationality-based homogeneity between the authors and reviewers. We defined national homogeneity as a condition with at least one member of the reviewer team (Reviewing Editor and peer reviewers) listing the same national affiliation as the corresponding author. We only considered the nationality of the corresponding author, since the nationality of the corresponding author was identical to the nationality of the first and last author for 94.1 and 94.5 percent of full submissions, respectively. Outside of the United States, the presence of country homogeneity during review was rare. While 90.5 percent of full submissions with corresponding authors from the U.S. were reviewed by at least one gatekeeper from their country, homogeneity was present for only 29.8 percent of full submissions with corresponding authors from the United Kingdom and 25.3 percent of those with a corresponding author from Germany. The likelihood of reviewer homogeneity falls sharply for Japan and China, which had author-reviewer homogeneity for only 7.6 and 6.4 percent of full submissions, respectively. More extensive details on the rate of author/reviewer homogeneity for each country can be found in S3 Table.

We examined whether author-reviewer homogeneity tended to result in the favoring of submissions from authors of the same country as the reviewer. We first pooled together all countries, as shown in Fig 7, and found that the presence of homogeneity during review was significantly associated with a higher accept rate, $\chi^2(1, n = 6,508) = 75.9$, $p < 0.0001$. However, most cases of homogeneity occurred for authors from the United States, so this result could potentially reflect the higher accept

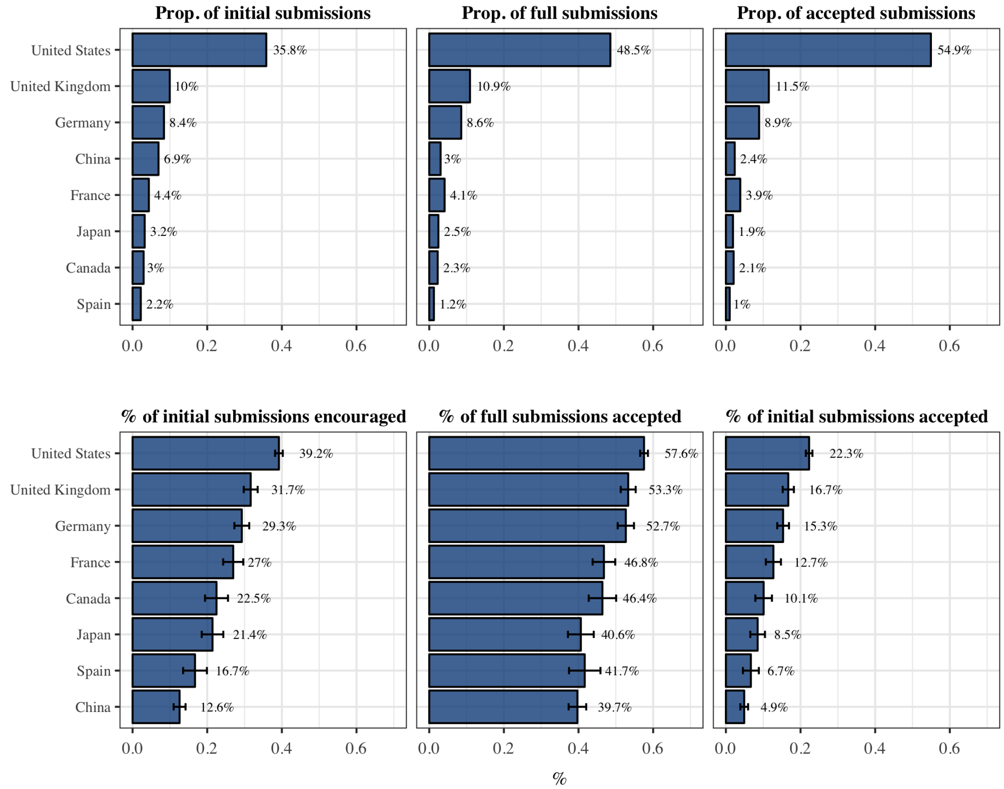


Fig 6. Share of submissions and outcomes of submissions by country of affiliation
Top: proportion of all initial submissions, encouraged initial submissions, and accepted full submissions comprised by the national affiliation of the corresponding author for the top eight most prolific countries in terms of initial submissions. Bottom: Encourage rate of initial submissions, acceptance rate of initial submissions, and acceptance rate of full submissions by national affiliation of the corresponding author for the top eight more prolific countries in terms of initial submissions. Error bars on bottom panel indicate standard error of proportion of encouraged initial submissions and accepted initial and full submissions for each country. This same graph with the top 16 most prolific nations can be found in S4 Fig.

rate for these authors, rather than homophily. Therefore we repeated the test, excluding all full submissions with corresponding authors from the United States, and we again found a significant homophilic effect, $\chi^2(1, n = 3,236) = 14.1, p < 0.001$. We repeated once more, excluding full submissions with corresponding authors from the the United States, United Kingdom, and Germany, and we identified no homophilic effect, $\chi^2(1, n = 1,920) = 0.095, p > 0.1$.

We also examined the effects of homogeneity within individual nations and tested for the presence of homophilic effects. Fig 7 shows accept rates for the eight most prolific nations submitting to *eLife*. For the United States, there was a weak relationship between the percentage of accepted full submissions and the presence of national homogeneity, $\chi^2(1, n = 3,270) = 2.9, p =< 0.1$. We observed a similar weak relationship for the United Kingdom, $\chi^2(1, n = 739) = 3.3, p < 0.1$. For China, we observed a statistically significant homophilic relationship between the acceptance rate of full submissions and national homogeneity $\chi^2(1, n = 204) = 5.2, p < 0.05$. We observed the

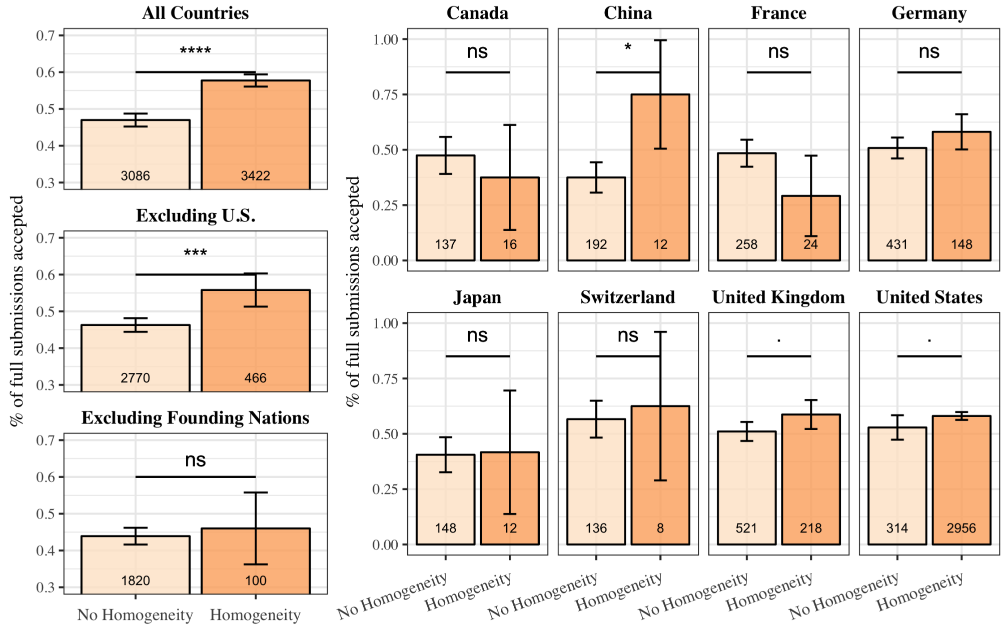


Fig 7. Interaction between author nationality and national-composition of gatekeepers.
Left: acceptance rate of full submissions compared between presence and absence of homogeneity between the national affiliation of the corresponding author and of at least one. Difference is shown comparing the results for all submissions (top), for all submissions that do not have corresponding authors from the U.S. (middle), and for all submissions that do not have a corresponding author from the U.S., U.K., or Germany (bottom). Right: acceptance rate of full submissions by national homogeneity, shown by individual countries. Included here were the top eight most prolific countries in terms of number of initial submissions. For both panels: vertical error bars indicate 95th percentile confidence intervals for the proportion of accepted full submissions. Values at the base of each bar indicate the number of observations within that combination of country and homophily variables. Asterisks indicate significance level of χ^2 tests of independence comparing frequency of accepted full submissions between presence and absence of homophily and within each country. "****" $= p < 0.0001$; "***" $= p < 0.001$; "* " $= p < 0.05$; "." $= p \leq 0.1$; "ns" $= p > 0.1$.

inverse trend for France and Canada, where the presence of gatekeepers from the same country was associated with lower accept rates, though this trend was not statistically significant. In summary, we found that the presence of national homogeneity was rare unless an author was from the United States, but that author/reviewer homogeneity was often (though not always) associated with homophilic bias.

Discussion

We identified inequities in peer review outcomes at *eLife*, based on gender and nationality of the senior (last and corresponding) authors. We observed a significant disparity in the acceptance rates of submissions with male and female last authors, which favored men. Inequities were also observed by country of affiliation. In particular, submissions from highly developed countries, with high scientific capacities, tended to have higher success rates than others. These inequities in peer review outcomes could be attributed, at least in part, to an interaction between gatekeeper and author demographics, which can be described as homophily, or a preference based on shared characteristics: Gatekeepers were more likely to recommend a manuscript for acceptance if they shared demographic characteristics with the authors. In particular, manuscripts with male senior (last or corresponding) authors were more likely to be accepted if reviewed by an all-male reviewer panel rather than a mixed-gender panel. Similarly, manuscripts were more likely to be accepted if at least one of the reviewers was from the same country as the last or corresponding author. The differential outcomes on the basis of homophily suggests that peer review at *eLife* is influenced by some form of bias—be it implicit bias [3, 16], geographic or linguistic bias [24, 69, 70], or cognitive particularism [40]. Specifically, a homophilic interaction suggests that peer review outcomes may sometimes be based on more than the intrinsic quality of manuscript; the composition of the review team is also related to outcomes in peer review.

The opportunity for homophilous interactions is determined by the demographics of the gatekeeper pool. We found that the demographics of the gatekeepers differed significantly from those of the authors, even for last authors, who tend to be more senior [65–68]. Women were underrepresented among *eLife* gatekeepers, and gatekeepers tended to come from a small number of highly-developed countries. The underrepresentation of women at *eLife* mirrors global trends—women comprise a minority of total authorships, yet constitute an even smaller proportion of gatekeepers across many domains [13, 71–78]. Similarly, gatekeepers at *eLife* were less internationally diverse than their authorship, reflecting the general underrepresentation of the "global south" in leadership positions of international journals [79].

The demographics of the reviewer pool made certain authors more likely to benefit from homophily in the review process than others. US authors were much more likely than not (see S3 Table) to be reviewed by a panel with at least one reviewer from the US. However, the opposite was true for authors from other countries. Fewer opportunities for such homophily may result in a disadvantage for scientists from smaller and less scientifically prolific countries. For gender, male lead authors had a nearly 50 percent chance of being reviewed by a homophilous (all-male), rather than a mixed-gender team. In contrast, because all-female reviewer panels were so rare (accounting for only 81 of 6,509 full submission decisions), female authors were highly unlikely to benefit from homophily in the review process.

Increasing *eLife*'s representation of women and scientists from a more diverse set of nations among editors may lead to more diverse reviewer pool and a more equitable peer review process. Editors often invite peer reviewers from their own professional networks, networks that likely reflect the characteristics of the editor [80–82]; this can lead to editors, who tend to be men [13, 71–78] and from scientifically advanced countries [79] to invite peer reviewers who are cognitively or demographically similar to themselves [44, 83, 84], inadvertently excluding certain groups from the gatekeeping process. Accordingly, we found that male Reviewing Editors at *eLife* were less likely to create mixed-gender teams of gatekeepers than female Reviewing Editors (see S5 Fig). We observed a similar effect based on the nationality of the Reviewing Editor and invited peer reviewers (see S6 Fig).

The size of disparities we observe in peer review outcomes may seem modest, however these small disparities can accumulate through each stage of the review process (initial submission, full submission, revisions), and potentially affect the outcomes of many submissions. For example, the overall acceptance rate (the rate at which initial submissions were eventually accepted) for male and female corresponding authors was 15.4 and 13.6 percent respectively; in other words, manuscripts submitted to *eLife* with female lead authors were published at 88.3 percent the rate of those with male lead authors. Similarly, manuscripts submitted by lead authors from China were accepted at only 22.0 percent the rate of manuscripts submitted by a lead author from the United States (with overall acceptance rates of 4.9 and 22.3 percent, respectively). Success in peer review is vital for a researcher's career because successful publication strengthens their professional reputation and makes it easier to attract funding, students, postdocs, and hence further publications. Even small advantages can compound over time and result in pronounced inequalities in science [85–88].

Our finding that the gender of the last authors is associated with a significant difference in the rate at which full submissions were accepted at *eLife* stands in contrast with a number of previous studies of journal peer review; these studies have found no significant difference in outcomes of papers submitted by male and female authors [55, 89, 90], or differences in reviewer's evaluations based on the author's apparent gender [91]. This discrepancy may may be explained in part by *eLife*'s unique context, policies, or the relative selectivity of *eLife* compared to venues where previous studies found gender equity. In addition, our results point to a key feature of study design that may account for some of the differences across studies, which is the consideration of multiple authorship roles. This is especially important for the biosciences, where authorship order is strongly associated with contribution [67, 68, 92]. Whereas our study examines the gender of the first, last, and corresponding authors, most previous studies have focused on the gender of the first author (e.g., [2, 89, 93]) or of the corresponding author (e.g., [21, 94]). Like previous studies, we observed no strong relationship between first author gender and review outcomes at *eLife*. Only when considering lead authorship roles—last authorship, and to a lesser extent, corresponding author, did we observe such an effect. Our results may be better compared with studies of grant peer review, where leadership roles are more explicitly defined, and many studies have identified significant disparities in outcomes favoring men [17, 18, 95–98], although many other studies have found no evidence of gender disparity [20, 22, 23, 99–101]. Given that science has grown increasingly collaborative and that average authorship per paper has expanded [102, 103], future studies of disparities would benefit from explicitly accounting for multiple authorship roles and signaling among various leadership positions on the byline [65, 104].

The interaction we found between the gender and nationality of the gatekeepers and peer review outcomes also stands in contrast to the findings from a number of previous studies. One study, [105], identified a homophilous relationship between female reviewers and female authors. However, most previous analyses found only procedural differences based on the gender of the gatekeeper [21, 90, 91, 106] and identified no difference in outcomes based on the interaction of author and gatekeeper gender in journal submissions [90, 107, 108] or grant review [22]. Studies of gatekeeper nationality have found no difference in peer review outcomes based on the nationality of the reviewer [107, 109], though there is little research on the correspondence between author and reviewer gender. One past study examined the interaction between U.S. and non-U.S. authors and gatekeepers, but found an effect opposite to what we observed, such that U.S. reviewers tended to rate submissions of U.S. authors more harshly than those of non-U.S. authors [43]. Our results also contrast with the study most similar to our own, which found no evidence of bias related to gender, and only modest evidence of bias related to geographic region [2]. These discrepancies may result from our analysis of multiple author roles. Alternatively, they may result from the unique nature of *eLife's* consultative peer review; the direct communication between peer reviewers compared to traditional peer review may render the social characteristics of reviewers more influential.

Limitations

There are limitations of our methodology that must be considered. First, we have no objective measure of the intrinsic quality of manuscripts. Therefore, it is not clear which review condition (homophilic or non-homophilic) more closely approximates the ideal of merit-based peer review outcomes. Second, measuring the interaction between reviewer and author demographics on peer review outcomes cannot readily detect biases that are shared by all reviewers/gatekeepers (e.g., if all reviewers, regardless of gender, favored manuscripts from male authors); hence, our approach could underestimate the influence of bias. Third, our analysis is observational, so we cannot establish causal relationships between success rates and authors or gatekeeper demographics. Along these lines, the reliance on statistical tests with arbitrary significance thresholds may provide misleading results (see [110]), or obfuscate statistically weak but potentially important relationships. Fourth, our gender-assignment algorithm is only a proxy for author gender and varies in reliability by continent.

Further studies will be required to determine the extent to which the effects we observe generalize to other peer review contexts. Specific policies at *eLife*, such as their consultative peer review process, may contribute to the effects we observed. Other characteristics of *eLife* may also be relevant, including its level of prestige [12], and its disciplinary specialization in the biological sciences, whose culture may differ from other scientific and academic disciplines. Future work is necessary to confirm and expand upon our findings, assess the extent to which they can be generalized, establish causal relationships, and mitigate the effects of these methodological limitations. To aid in this effort, we have made as much as possible of the data and analysis publicly available at (https://github.com/murrayds/elife-analysis).

Conclusion and recommendations

Many factors can contribute to gender, national, and other inequities in scientific publishing. This includes a variety of factors entirely external to peer review [46, 50, 111–114], which can affect the quantity and perceived quality of submitted manuscripts. However, these structural factors do not readily account for the observed interaction between gatekeeper and author demographics associated with peer review outcomes at *eLife*; rather, biases related to the personal characteristics of the authors and gatekeepers are likely to play some role in peer review outcomes.

Our results suggest that it is not only the form of peer review that matters, but also the composition of reviewers. Homophilous preferences in evaluation are a potential mechanism underpinning the Matthew Effect [1] in academia. This effect entrenches privileged groups while potentially limiting diversity, which could hinder scientific production, since diversity may lead to better working groups [115] and promote high-quality science [116, 117]. Increasing gender and international representation among scientific gatekeepers may improve fairness and equity in peer review outcomes, and accelerate scientific progress. However, this must be carefully balanced to avoid overburdening scholars from minority groups with disproportionate service obligations.

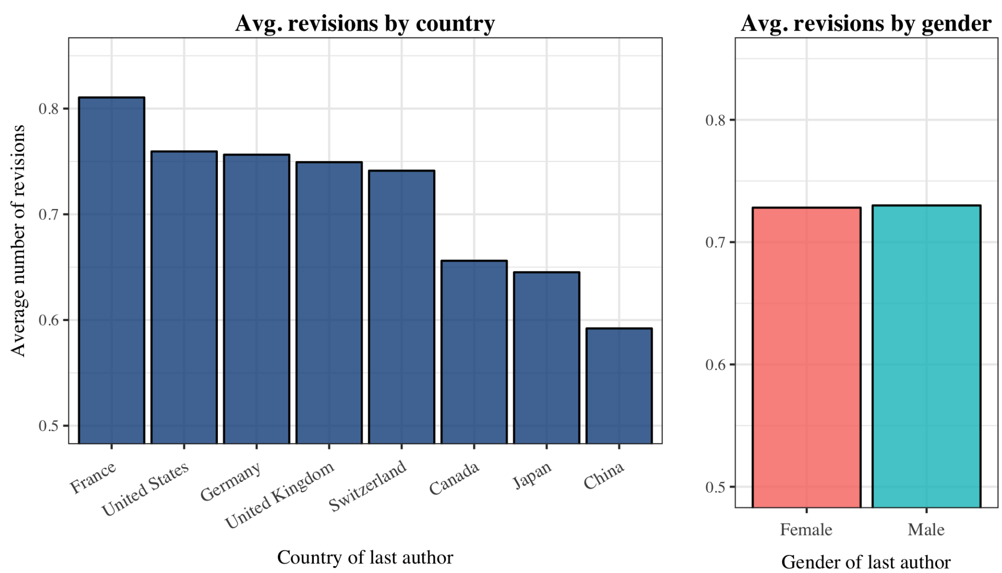
Although some journals, such as *eLife* and Frontiers Media, have begun providing peer review data to researchers (see [44, 118]), data on equity in peer review outcomes is currently available only for a small fraction of journals and funders. While many journals collect these data internally, they are not usually standardized or shared publicly. One group, PEERE, authored a protocol for open sharing of peer review data [119, 120], though this protocol is recent, and the extent to which it will be adopted remains uncertain. To both provide better benchmarks and to incentivize better practices, journals should make analyses on author and reviewer demographics publicly available. These data include, but would not be limited to, characteristics such as gender, race, sexual orientation, seniority, and institution and country of affiliation. It is likely that privacy concerns and issues relating to confidentiality will limit the full availability of the data; but analyses that are sensitive to the vulnerabilities of smaller populations should be conducted and made available as benchmarking data.

Some high-profile journals have experimented with implementing double-blind peer review as a potential solution to inequities in publishing, including *Nature* [121] and *eNeuro* [11], though in some cases with mixed results [60]. In addition, journals are analyzing the demographics of their published authorship and editorial staff in order to identify key problem areas, focus initiatives, and track progress in achieving diversity goals [13, 83, 89]. Alternatives to traditional peer review have also been proposed, including open peer review, study pre-registration, consultative peer review, and hybrid processes (eg: [64, 122–126]), as well as alternative forms of dissemination, such as preprint servers (e.g., arXiv, bioRxiv). Currently, there is little empirical evidence to determine whether these formats constitute less biased or more equitable alternatives [3].

More work should be done to study and understand the issues facing peer review and scientific gatekeeping in all its forms, and to promote fair, efficient, and meritocratic scientific cultures and practices. Editorial bodies should craft policies and implement practices that diminish disparities in peer review; they should also continue to be innovative and reflective about their practices to ensure that papers are accepted on scientific merit, rather than particularistic characteristics of the authors.

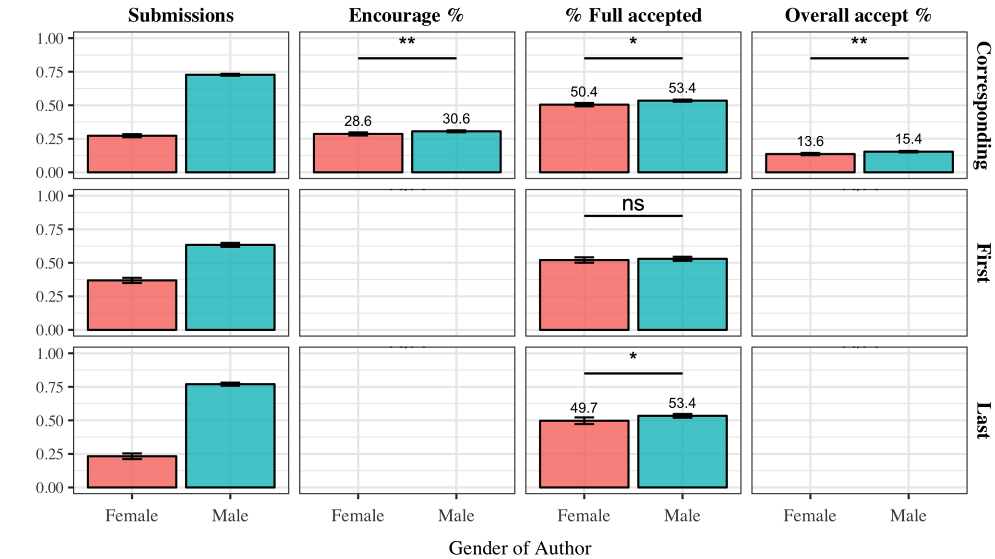
Supporting information

S1 Fig.



Number of revisions by author gender and nationality. Average number of revisions a full submissions undergoes before a final decision of accept or reject is made. In this case, zero revisions occurs when a full submission is accepted or rejected without a request for any revisions. The dataset records at maximum two revisions, though only a small number of manuscripts remain in revision after two submissions (see Fig 1). For this figure, we only include manuscripts for which a final decision is made after zero, one, or two revisions. The left panel shows differences in the average number of revisions by the country of the last author. The right shows the average revisions by the gender of the last author.

S2 Fig.

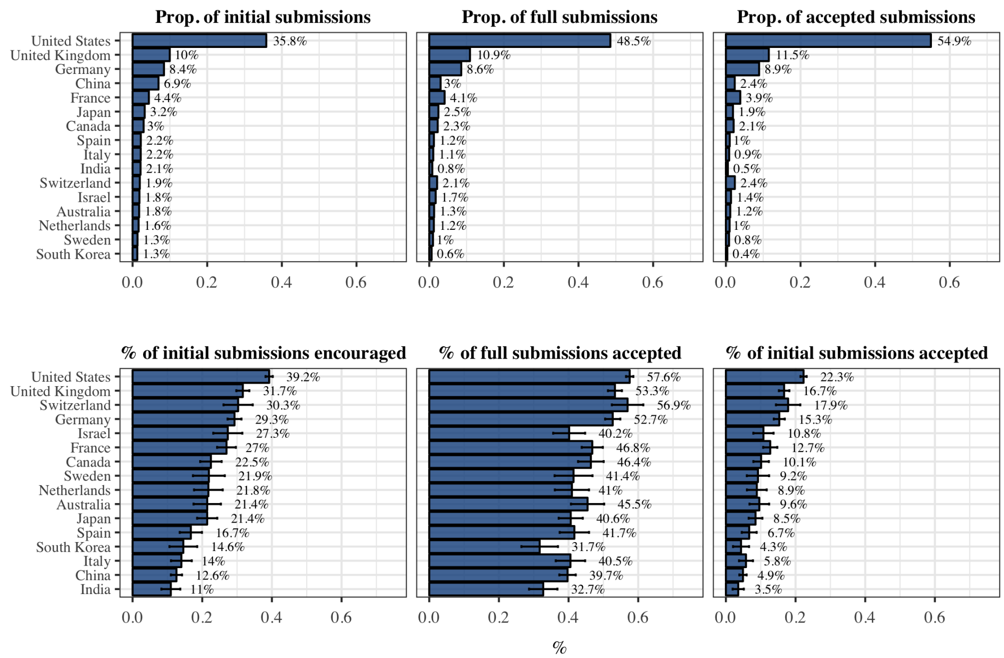


Submission and success rates by gender of corresponding, first, and last author. Proportion of initial submissions, encourage rate, overall accept rate, and accept rate of full submissions by the gender of the corresponding author, first author, and last author. Gender data is unavailable for first and last authors of initial submissions that were never submitted as full submissions, therefore these cells remain blank. Authors whose gender is unknown are excluded from analysis. Vertical error bars indicate 95$^{\text{th}}$ confidence intervals of the proportion of submitted, encouraged, and accepted initial and full submissions. Asterisks indicate significance level of χ^2 tests of independence of frequency of encourage and acceptance by gender; "*" $= p < 0.05$; "ns" $= p > 0.1$.

S3 Fig.

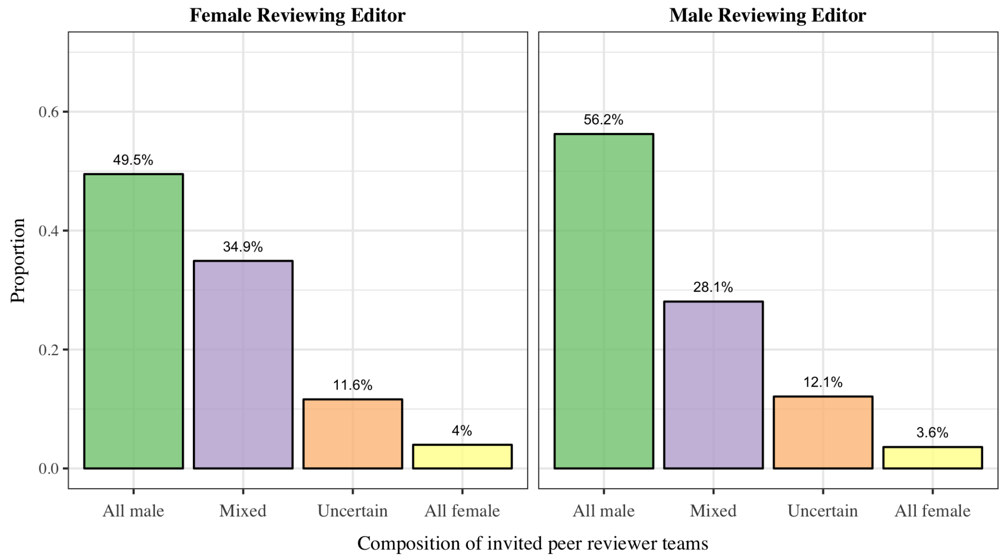
Submission and success rates by gender of corresponding, first, and last author. Percentage of full submissions that were accepted, shown by the gender of the corresponding, first, and last author, and by the gender composition of the peer reviewers. Text at the base of each bar indicate the number full submissions within each category of reviewer team and authorship gender. Vertical error bars indicate 95$^{\text{th}}$ percentile confidence intervals of the proportion of accepted full submissions. Asterisks indicate significance level of chi^2 tests of independence on frequency of acceptance by gender of author given each team composition. "*" $= p < 0.05$; "ns" $= p > 0.1$.

S4 Fig.



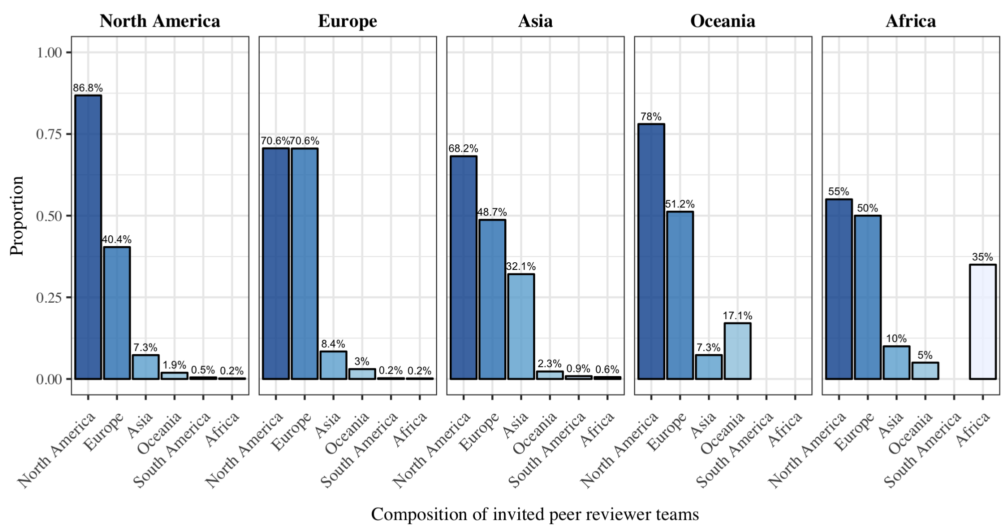
Submission and success rates by country for top 16 most prolific countries. Top: proportion of all initial submissions, encouraged initial submissions, and accepted full submissions comprised by the national affiliation of the corresponding author for the top sixteen most prolific countries in terms of initial submissions. Bottom: acceptance rate of full submissions, encourage rate of full submissions, and overall accept rate of full submissions by national affiliation of the corresponding author for the top eight more prolific countries in terms of initial submissions. Error bars on bottom panel indicate standard error of proportion of encouraged initial submissions and accepted initial and full submissions for each country.

S5 Fig.



Proportion of peer reviewer team's gender compositions by gender of the Reviewing Editor. Compositions are determined while excluding the Reviewing Editor from team membership, if they are listed as a peer reviewer.

S6 Fig.



Proportion of peer review teams containing at least one peer reviewer of each continent, by continent of Reviewing Editor. Compositions are determined while excluding the Reviewing Editor from team membership, if they are listed as a peer reviewer.

S1 Table. Summary demographic characteristics of distinct *eLife* reviewers and editors. The count of Senior Editors includes former editors, as well as the Deputy Editors and Editor-in-Chief, who also serve as Senior Editors. The count of BREs includes former editors and guest editors. Reviewers are only relevant for publications that were submitted for full review, thus leading to lower total counts. Includes all individuals involved in processing manuscripts at *eLife* between 2012 and 2017.

Reviewership	**Female**		**Male**		**Unassigned**		
	N	**%**	**N**	**%**	**N**	**%**	**All**
Senior Editors	15	26.3	42	73.7	0	0.0	57
Reviewing Editors	209	24.0	661	76.0	0.0	0.0	870
Peer Reviewers	1,439	20.7	5,266	75.7	256	3.60	6,961

S2 Table. Summary nationality demographics of unique *eLife* reviewers and editors. The count of Senior Editors includes former editors, as well as the Deputy Editors and Editor-in-Chief, who also serve as Senior Editors. The count of BREs includes former editors and guest editors. Reviewers are only relevant for publications that were submitted for full review, thus leading to lower total counts. Includes all individuals involved in processing manuscripts at *eLife* between 2012 and 2017.

Country	Senior Editor		Reviewing Editor		Peer Reviewer	
	N	%	N	%	N	%
United States	32	56.1	536	62	3648	55.9
United Kingdom	7	12.3	88	10.2	707	10.8
Germany	6	10.5	69	8	469	7.2
Canada	3	5.3	22	2.5	249	3.8
Switzerland	2	3.5	19	2.2	166	2.5
China	2	3.5	10	1.2	61	0.9
Israel	1	1.8	19	2.2	89	1.4
Netherlands	1	1.8	11	1.3	114	1.7
Spain	1	1.8	10	1.2	78	1.2
Japan	1	1.8	9	1	128	2
India	1	1.8	6	0.7	20	0.3
France	0	0	21	2.4	245	3.8
Australia	0	0	7	0.8	87	1.3
South Africa	0	0	5	0.6	12	0.2
Austria	0	0	4	0.5	49	0.8
Belgium	0	0	3	0.3	37	0.6

S3 Table. Submissions and proportion of author/gatekeeper homogeneity by country. Includes number of full submissions submitted with corresponding authors from each of 20 countries, and proportion of these full submissions with the condition of author/reviewer homogeneity such that at least one involved gatekeeper from the same country. Countries listed are in order of the proportion of author/reviewer homogeneity, and contain the top 20 countries with the highest homogeneity.

Corr. Author Country	# Submissions	Proportion Homogeneity
United States	3,605	0.91
United Kingdom	803	0.3
Germany	641	0.25
Canada	176	0.11
South Korea	45	0.11
South Africa	11	0.09
France	310	0.08
Japan	184	0.08
Australia	101	0.07
China	233	0.06
Switzerland	163	0.06
India	59	0.05
Sweden	70	0.04
Israel	127	0.04
Spain	91	0.03
Denmark	32	0.03
Italy	79	0.03
Belgium	41	0.02
Austria	58	0.02
Netherlands	100	0.01

Acknowledgments

We are grateful for the editing and feedback provided by Susanna Richmond (Senior Editorial Assistant at *eLife*), Mark Patterson (Executive Director at *eLife*), Eve Marder, Anna Akhmanova, and Detlef Weigel (Deputy Editors at *eLife*). We are also grateful for the work of James Gilbert (Production Editor at *eLife*) for extracting the data used in this analysis. This work was partially supported by a grant from the National Science Foundation (SciSIP #1561299).

Competing interests

Wei Mun Chan and Andrew M. Collings are employed by *eLife*. Jennifer Raymond and Cassidy R. Sugimoto are Reviewing Editors at *eLife*. Andrew M. Collings was employed by PLOS between 2005 and 2012.

References

1. Merton RK. The Matthew Effect in Science: The reward and communication systems of science are considered. Science. 1968;159(3810):56–63. doi:10.1126/science.159.3810.56.

2. Walker R, Barros B, Conejo R, Neumann K, Telefont M. Personal attributes of authors and reviewers, social bias and the outcomes of peer review: a case study. F1000Research. 2015;doi:10.12688/f1000research.6012.2.

3. Lee CJ, Sugimoto CR, Zhang G, Cronin B. Bias in peer review. Journal of the American Society for Information Science and Technology. 2013;64(1):2–17. doi:10.1002/asi.22784.

4. Pinholster G. Journals and funders confront implicit bias in peer review. Science. 2016;352(6289):1067–1068. doi:10.1126/science.352.6289.1067.

5. Larivière V, Ni C, Gingras Y, Cronin B, Sugimoto CR. Bibliometrics: Global gender disparities in science. Nature News. 2013;504(7479):211. doi:10.1038/504211a.

6. Women, Minorities, and Persons with Disabilities in Science and Engineering: 2017. Arlington, VA: National Science Foundation, National Center for Science and Engineering Statistics.; 2017.

7. Gender in the Global Research Landscape. Elsevier; 2017. Available from: `https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf`.

8. West JD, Jacquet J, King MM, Correll SJ, Bergstrom CT. The Role of Gender in Scholarly Authorship. PLOS ONE. 2013;8(7):e66212. doi:10.1371/journal.pone.0066212.

9. Larivière V, Sugimoto CR. The end of gender disparities in science? If only it were true...; 2017. Available from: `https://www.cwts.nl:443/blog?article=n-q2z294&title=the-end-of-gender-disparities-in-science-if-only-it-were-true`.

10. Bendels MHK, Müller R, Brueggmann D, Groneberg DA. Gender disparities in high-quality research revealed by Nature Index journals. PLOS ONE. 2018;13(1):e0189136. doi:10.1371/journal.pone.0189136.

11. Bernard C. Editorial: Gender Bias in Publishing: Double-Blind Reviewing as a Solution? eNeuro. 2018;5(3):ENEURO.0225–18.2018. doi:10.1523/ENEURO.0225-18.2018.

12. Shen YA, Webster JM, Shoda Y, Fine I. Persistent Underrepresentation of Women's Science in High Profile Journals. bioRxiv. 2018; p. 275362. doi:10.1101/275362.

13. Nature's under-representation of women. Nature. 2018;558(7710):344. doi:10.1038/d41586-018-05465-7.

14. King DA. The scientific impact of nations. Nature. 2004;430:311.

15. White KE, Robbins C, Khan B, Freyman C. Science and Engineering Publication Output Trends: 2014 Shows Rise of Developing Country Output while Developed Countries Dominate Highly Cited Publications. Arlington, VA: National Center for Science and Engineering Statistics; 2017. NSF 18-300.

16. Wennerås C, Wold A. Nepotism and sexism in peer-review. Nature. 1997;387:341.
17. Li D. Gender Bias in NIH Peer Review: Does it Exist and Does it Matter?; 2011. Available from: `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.571.8269&rep=rep1&type=pdf`.
18. Tamblyn R, Girard N, Qian CJ, Hanley J. Assessment of potential bias in research grant peer review in Canada. CMAJ. 2018;190(16):E489–E499. doi:10.1503/cmaj.170901.
19. Witteman HO, Hendricks M, Straus S, Tannenbaum C. Female grant applicants are equally successful when peer reviewers assess the science, but not when they assess the scientist. bioRxiv. 2017; p. 232868. doi:10.1101/232868.
20. Grant J, Burden S, Breen G. No evidence of sexism in peer review. Nature. 1997;390(6659):438. doi:10.1038/37213.
21. Gilbert JR, Williams ES, Lundberg GD. Is there gender bias in JAMA's peer review process? JAMA. 1994;272(2):139–142.
22. Mutz R, Bornmann L, Daniel HD. Does Gender Matter in Grant Peer Review?: An Empirical Investigation Using the Example of the Austrian Science Fund. Zeitschrift Fur Psychologie. 2012;220(2):121–129. doi:10.1027/2151-2604/a000103.
23. Beck R, Halloin V. Gender and research funding success: Case of the Belgian F.R.S.-FNRS. Research Evaluation. 2017;26(2):115–123. doi:10.1093/reseval/rvx008.
24. Coates R, Sturgeon B, Bohannan J, Pasini E. Language and publication in Cardiovascular Research articles. Cardiovascular Research. 2002;53(2):279–285. doi:10.1016/S0008-6363(01)00530-2.
25. Primack RB, Marrs R. Bias in the review process. Biological Conservation. 2008;141(12):2919–2920. doi:10.1016/j.biocon.2008.09.016.
26. Harris M, Marti J, Watt H, Bhatti Y, Macinko J, Darzi AW. Explicit Bias Toward High-Income-Country Research: A Randomized, Blinded, Crossover Experiment Of English Clinicians. Health Affairs (Project Hope). 2017;36(11):1997–2004. doi:10.1377/hlthaff.2017.0773.
27. Primack RB, Ellwood E, Miller-Rushing AJ, Marrs R, Mulligan A. Do gender, nationality, or academic age affect review decisions? An analysis of submissions to the journal Biological Conservation. Biological Conservation. 2009;142(11):2415–2418. doi:10.1016/j.biocon.2009.06.021.
28. Yousefi-Nooraie R, Shakiba B, Mortaz-Hejri S. Country development and manuscript selection bias: a review of published studies. BMC Medical Research Methodology. 2006;6:37. doi:10.1186/1471-2288-6-37.
29. Berger J, Fisek HF, Normal RZ, Zelditch N. Status characteristics and social interaction. New York: Elsevier; 1977.
30. Correll SJ, Ridgeway CL. Expectation States Theory. In: Handbook of Social Psychology. Handbooks of Sociology and Social Research. Springer, Boston, MA; 2006. p. 29–51. Available from: `https://link.springer.com/chapter/10.1007/0-387-36921-X_2`.

31. Budden AE, Tregenza T, Aarssen LW, Koricheva J, Leimu R, Lortie CJ. Double-blind review favours increased representation of female authors. Trends in Ecology & Evolution. 2008;23(1):4–6. doi:10.1016/j.tree.2007.07.008.

32. Podolny JM. Status Signals: A Sociological Study of Market Competition. Princeton, N.J Woodstock: Princeton University Press; 2008.

33. Long JS, Fox MF. Scientific Careers: Universalism and Particularism. Annual Review of Sociology. 1995;21:45–71.

34. Pfeffer J, Leong A, Strehl K. Paradigm Development and Particularism: Journal Publication in Three Scientific Disciplines. Social Forces. 1977;55(4):938–951. doi:10.2307/2577563.

35. Cole S. The Hierarchy of the Sciences? American Journal of Sociology. 1983;89(1):111–139. doi:10.1086/227835.

36. Clauset A, Arbesman S, Larremore DB. Systematic inequality and hierarchy in faculty hiring networks. Science Advances. 2015;1(1):e1400005. doi:10.1126/sciadv.1400005.

37. Jacobs JA. Gender and the Stratification of Colleges. The Journal of Higher Education. 1999;70(2):161–187. doi:10.2307/2649126.

38. Weeden KA, Thebaud S, Gelbgiser D. Degrees of Difference: Gender Segregation of U.S. Doctorates by Field and Program Prestige. Sociological Science. 2017;4:123–150. doi:10.15195/v4.a6.

39. Sheltzer JM, Smith JC. Elite male faculty in the life sciences employ fewer women. Proceedings of the National Academy of Sciences. 2014;111(28):10107–10112. doi:10.1073/pnas.1403334111.

40. Travis GDL, Collins HM. New Light on Old Boys: Cognitive and Institutional Particularism in the Peer Review System. Science, Technology, & Human Values. 1991;16(3):322–341.

41. Teplitskiy M, Acuna D, Elamrani-Raoult A, Körding K, Evans J. The sociology of scientific validity: How professional networks shape judgement in peer review. Research Policy. 2018;doi:10.1016/j.respol.2018.06.014.

42. Demarest B, Freeman G, Sugimoto CR. The reviewer in the mirror: examining gendered and ethnicized notions of reciprocity in peer review. Scientometrics. 2014;101(1):717–735. doi:10.1007/s11192-014-1354-z.

43. Link AM. US and non-US submissions: an analysis of reviewer bias. JAMA. 1998;280(3):246–247.

44. Helmer M, Schottdorf M, Neef A, Battaglia D. Research: Gender bias in scholarly peer review. eLife. 2017;6:e21718. doi:10.7554/eLife.21718.

45. Bagues M, Sylos-Labini M, Zinovyeva N. Does the Gender Composition of Scientific Committees Matter? American Economic Review. 2017;107(4):1207–1238. doi:10.1257/aer.20151211.

46. Ceci SJ, Williams WM. Understanding current causes of women's underrepresentation in science. Proceedings of the National Academy of Sciences. 2011;108(8):3157–3162. doi:10.1073/pnas.1014871108.

47. Ceci SJ, Ginther DK, Kahn S, Williams WM. Women in Academic Science: A Changing Landscape. Psychological Science in the Public Interest: A Journal of the American Psychological Society. 2014;15(3):75–141. doi:10.1177/1529100614541236.

48. Niederle M, Vesterlund L. Gender and Competition. Annual Review of Economics. 2011;3(1):601–630. doi:10.1146/annurev-economics-111809-125122.

49. Hengel E. Publishing while Female. Are women held to higher standards? Evidence from peer review. Faculty of Economics, University of Cambridge; 2017. 1753. Available from: `https://ideas.repec.org/p/cam/camdae/1753.html`.

50. May RM. The Scientific Wealth of Nations. Science. 1997;275(5301):793–796. doi:10.1126/science.275.5301.793.

51. Quan W, Chen B, Shu F. Publish or impoverish: An investigation of the monetary reward system of science in China (1999-2016). Aslib Journal of Information Management. 2017;69(5):486–502. doi:10.1108/AJIM-01-2017-0014.

52. Duszak A, Lewkowicz J. Publishing academic texts in English: A Polish perspective. Journal of English for Academic Purposes. 2008;7(2):108–120. doi:10.1016/j.jeap.2008.03.001.

53. Salager-Meyer F. Scientific publishing in developing countries: Challenges for the future. Journal of English for Academic Purposes. 2008;7(2):121–132. doi:10.1016/j.jeap.2008.03.009.

54. Yang W. Policy: Boost basic research in China. Nature News. 2016;534(7608):467. doi:10.1038/534467a.

55. Tomkins A, Zhang M, Heavlin WD. Reviewer bias in single- versus double-blind peer review. Proceedings of the National Academy of Sciences. 2017; p. 201707323. doi:10.1073/pnas.1707323114.

56. Blank RM. The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence from The American Economic Review. The American Economic Review. 1991;81(5):1041–1067.

57. Okike K, Hug KT, Kocher MS, Leopold SS. Single-blind vs Double-blind Peer Review in the Setting of Author Prestige. JAMA. 2016;316(12):1315–1316. doi:10.1001/jama.2016.11014.

58. Ware M. Peer Review in Scholarly Journals: Perspective of the Scholarly Community - Results from an International Study. Inf Serv Use. 2008;28(2):109–112.

59. Kmietowicz Z. Double blind peer reviews are fairer and more objective, say academics. BMJ : British Medical Journal. 2008;336(7638):241. doi:10.1136/bmj.39476.357280.DB.

60. McGillivray B, De Ranieri E. Uptake and outcome of manuscripts in Nature journals by review model and author characteristics. arXiv:180202188 [cs]. 2018;.

61. Langfeldt L. The Decision-Making Constraints and Processes of Grant Peer Review, and Their Effects on the Review Outcome. Social Studies of Science. 2001;31(6):820–841. doi:10.1177/030631201031006002.

62. Lamont M. How Professors Think: Inside the Curious World of Academic Judgment. Cambridge: Harvard University Press; 2009.

63. Schekman R, Watt F, Weigel D. Scientific Publishing: The eLife approach to peer review. eLife. 2013;2:e00799. doi:10.7554/eLife.00799.

64. King SR. Peer Review: Consultative review is worth the wait. eLife. 2017;6:e32012. doi:10.7554/eLife.32012.

65. Costas R, Bordons M. Do age and professional rank influence the order of authorship in scientific publications? Some evidence from a micro-level perspective. Scientometrics. 2011;88(1):145–161. doi:10.1007/s11192-011-0368-z.

66. Macaluso B, Larivière V, Sugimoto T, Sugimoto CR. Is Science Built on the Shoulders of Women? A Study of Gender Differences in Contributorship. Academic Medicine: Journal of the Association of American Medical Colleges. 2016;91(8):1136–1142. doi:10.1097/ACM.0000000000001261.

67. Baerlocher MO, Newton M, Gautam T, Tomlinson G, Detsky AS. The meaning of author order in medical research. Journal of Investigative Medicine: The Official Publication of the American Federation for Clinical Research. 2007;55(4):174–180. doi:10.2310/6650.2007.06044.

68. Tscharntke T, Hochberg ME, Rand TA, Resh VH, Krauss J. Author Sequence and Credit for Contributions in Multiauthored Publications. PLOS Biology. 2007;5(1):e18. doi:10.1371/journal.pbio.0050018.

69. Kliewer MA, DeLong DM, Freed K, Jenkins CB, Paulson EK, Provenzale JM. Peer review at the American Journal of Roentgenology: how reviewer and manuscript characteristics affected editorial decisions on 196 major papers. AJR American journal of roentgenology. 2004;183(6):1545–1550. doi:10.2214/ajr.183.6.01831545.

70. Ross JS, Gross CP, Desai MM, Hong Y, Grant AO, Daniels SR, et al. Effect of blinded peer review on abstract acceptance. JAMA. 2006;295(14):1675–1680. doi:10.1001/jama.295.14.1675.

71. Amrein K, Langmann A, Fahrleitner-Pammer A, Pieber TR, Zollner-Schwetz I. Women Underrepresented on Editorial Boards of 60 Major Medical Journals. Gender Medicine. 2011;8(6):378–387. doi:10.1016/j.genm.2011.10.007.

72. Cho AH, Johnson SA, Schuman CE, Adler JM, Gonzalez O, Graves SJ, et al. Women are underrepresented on the editorial boards of journals in environmental biology and natural resource management. PeerJ. 2014;2:e542. doi:10.7717/peerj.542.

73. Mauleón E, Hillán L, Moreno L, Gómez I, Bordons M. Assessing gender balance among journal authors and editorial board members. Scientometrics. 2013;95(1):87–114. doi:10.1007/s11192-012-0824-4.

74. Metz I, Harzing AW. Gender Diversity in Editorial Boards of Management Journals. Academy of Management Learning & Education. 2009;8(4):540–557. doi:10.5465/amle.8.4.zqr540.

75. Metz I, Harzing AW. An update of gender diversity in editorial boards: a longitudinal study of management journals. Personnel Review. 2012;41(3):283–300. doi:10.1108/00483481211212940.

76. Morton MJ, Sonnad SS. Women on professional society and journal editorial boards. Journal of the National Medical Association. 2007;99(7):764–771.

77. Stegmaier M, Palmer B, Assendelft Lv. Getting on the Board: The Presence of Women in Political Science Journal Editorial Positions. PS: Political Science & Politics. 2011;44(4):799–804. doi:10.1017/S1049096511001284.

78. Topaz CM, Sen S. Gender Representation on Journal Editorial Boards in the Mathematical Sciences. PLOS ONE. 2016;11(8):e0161357. doi:10.1371/journal.pone.0161357.

79. Espin J, Palmas S, Carrasco-Rueda F, Riemer K, Allen PE, Berkebile N, et al. A persistent lack of international representation on editorial boards in environmental biology. PLOS Biology. 2017;15(12):e2002760. doi:10.1371/journal.pbio.2002760.

80. Addis E, Villa P. The Editorial Boards of Italian Economics Journals: Women, Gender, and Social Networking. Feminist Economics. 2003;9(1):75–91. doi:10.1080/1354570032000057062.

81. Avin C, Keller B, Lotker Z, Mathieu C, Peleg D, Pignolet YA. Homophily and the Glass Ceiling Effect in Social Networks. In: Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science. ITCS '15. New York, NY, USA: ACM; 2015. p. 41–50. Available from: `http://doi.acm.org/10.1145/2688073.2688097`.

82. Szell M, Thurner S. How women organize social networks different from men. Scientific Reports. 2013;3:1214. doi:10.1038/srep01214.

83. Lerback J, Hanson B. Journals invite too few women to referee. Nature News. 2017;541(7638):455. doi:10.1038/541455a.

84. Tite L, Schroter S. Why do peer reviewers decline to review? A survey. Journal of Epidemiology & Community Health. 2007;61(1):9–12. doi:10.1136/jech.2006.049817.

85. Bol T, Vaan Md, Rijt Avd. The Matthew effect in science funding. Proceedings of the National Academy of Sciences. 2018; p. 201719557. doi:10.1073/pnas.1719557115.

86. O'Brien KR, Hapgood KP. The academic jungle: ecosystem modelling reveals why women are driven out of research. Oikos. 2012;121(7):999–1004. doi:10.1111/j.1600-0706.2012.20601.x.

87. Petersen AM, Penner O. Inequality and cumulative advantage in science careers: a case study of high-impact journals. EPJ Data Science. 2014;3(1):24. doi:10.1140/epjds/s13688-014-0024-y.

88. Day TE. The big consequences of small biases: A simulation of peer review. Research Policy. 2015;44(6):1266–1270. doi:10.1016/j.respol.2015.01.006.

89. Valkonen L, Brooks J. Gender balance in Cortex acceptance rates. Cortex. 2011;47(7):763–770. doi:10.1016/j.cortex.2011.04.004.

90. Fox CW, Burns CS, Meyer JA, Thompson K. Editor and reviewer gender influence the peer review process but not peer review outcomes at an ecology journal. Functional Ecology. 2015;30(1):140–153. doi:10.1111/1365-2435.12529.

91. Borsuk RM, Aarssen LW, Budden AE, Koricheva J, Leimu R, Tregenza T, et al. To Name or Not to Name: The Effect of Changing Author Gender on Peer Review. BioScience. 2009;59(11):985–989. doi:10.1525/bio.2009.59.11.10.

92. Kassis T. How do research faculty in the biosciences evaluate paper authorship criteria? PLOS ONE. 2017;12(8):e0183632. doi:10.1371/journal.pone.0183632.

93. Tregenza T. Gender bias in the refereeing process? Trends in Ecology & Evolution. 2002;17(8):349–350. doi:10.1016/S0169-5347(02)02545-4.

94. Women in neuroscience: a numbers game. Nature Neuroscience. 2006;9(7):853. doi:10.1038/nn0706-853.

95. Bedi G, Van Dam NT, Munafo M. Gender inequality in awarded research grants. Lancet (London, England). 2012;380(9840):474. doi:10.1016/S0140-6736(12)61292-6.

96. Gannon F, Quirk S, Guest S. Searching for discrimination: Are women treated fairly in the EMBO postdoctoral fellowship scheme? EMBO reports. 2001;2(8):655–657. doi:10.1093/embo-reports/kve170.

97. Lee Rvd, Ellemers N. Gender contributes to personal research funding success in The Netherlands. Proceedings of the National Academy of Sciences. 2015;112(40):12349–12353. doi:10.1073/pnas.1510159112.

98. Shen H. Inequality quantified: Mind the gender gap. Nature. 2013;495(7439):22–24. doi:10.1038/495022a.

99. Pohlhaus JR, Jiang H, Wagner RM, Schaffer WT, Pinn VW. Sex Differences in Application, Success, and Funding Rates for Nih Extramural Programs. Academic Medicine. 2011;86(6):759–767. doi:10.1097/ACM.0b013e31821836ff.

100. Waisbren SE, Bowles H, Hasan T, Zou KH, Emans SJ, Goldberg C, et al. Gender differences in research grant applications and funding outcomes for medical school faculty. Journal of Women's Health (2002). 2008;17(2):207–214. doi:10.1089/jwh.2007.0412.

101. Marsh HW, Jayasinghe UW, Bond NW. Gender differences in peer reviews of grant applications: A substantive-methodological synergy in support of the null hypothesis model. Journal of Informetrics. 2011;5(1):167–180. doi:10.1016/j.joi.2010.10.004.

102. Lariviere V, Sugimoto C, Tsou A, Gingras Y. Team size matters: Collaboration and scientific impact since 1900. arXiv:14108544 [cs]. 2014;.

103. Wuchty S, Jones BF, Uzzi B. The Increasing Dominance of Teams in Production of Knowledge. Science. 2007;doi:10.1126/science.1136099.

104. Larivière V, Desrochers N, Macaluso B, Mongeon P, Paul-Hus A, Sugimoto CR. Contributorship and division of labor in knowledge production. Social Studies of Science. 2016;46(3):417–435. doi:10.1177/0306312716650046.

105. Lloyd ME. Gender factors in reviewer recommendations for manuscript publication. Journal of Applied Behavior Analysis. 1990;23(4):539–543. doi:10.1901/jaba.1990.23-539.

106. Wing DA, Benner RS, Petersen R, Newcomb R, Scott JR. Differences in editorial board reviewer behavior based on gender. Journal of Women's Health (2002). 2010;19(10):1919–1923. doi:10.1089/jwh.2009.1904.

107. Bornmann L, Daniel HD. Gatekeepers of science—Effects of external reviewers' attributes on the assessments of fellowship applications. Journal of Informetrics. 2007;1(1):83–91. doi:10.1016/j.joi.2006.09.005.

108. Petty RE, Fleming MA, Fabrigar LR. The Review Process at PSPB: Correlates of Interreviewer Agreement and Manuscript Acceptance. Personality and Social Psychology Bulletin. 1999;25(2):188–203. doi:10.1177/0146167299025002005.

109. Zhang X. Effect of reviewer's origin on peer review: China vs. non-China. Learned Publishing. 2012;25(4):265–270. doi:10.1087/20120405.

110. Gelman A, Stern H. The Difference Between "Significant" and "Not Significant" is not Itself Statistically Significant. The American Statistician. 2006;60(4):328–331. doi:10.1198/000313006X152649.

111. Adamo SA. Attrition of Women in the Biological Sciences: Workload, Motherhood, and Other Explanations Revisited. BioScience. 2013;63(1):43–48. doi:10.1525/bio.2013.63.1.9.

112. Ceci SJ, Williams WM, Barnett SM. Women's underrepresentation in science: sociocultural and biological considerations. Psychological Bulletin. 2009;135(2):218–261. doi:10.1037/a0014412.

113. Ceci SJ, Williams WM. Sex Differences in Math-Intensive Fields. Current Directions in Psychological Science. 2010;19(5):275–279. doi:10.1177/0963721410383241.

114. Xie Y, Shauman KA. Sex Differences in Research Productivity: New Evidence about an Old Puzzle. American Sociological Review. 1998;63(6):847–870. doi:10.2307/2657505.

115. Page SE. The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies. Princeton: Princeton University Press; 2008.

116. Campbell LG, Mehtani S, Dozier ME, Rinehart J. Gender-Heterogeneous Working Groups Produce Higher Quality Science. PLOS ONE. 2013;8(10):e79147. doi:10.1371/journal.pone.0079147.

117. Science & Engineering Indicators 2018. Alexandria, Virginia, United States: National Center for Science and Engineering Statistics (NCSES); 2018. NSB-2018-1.

118. Giordan M, Csikasz-Nagy A, Collings AM, Vaggi F. The effects of an editor serving as one of the reviewers during the peer-review process. F1000Research. 2016;5:683. doi:10.12688/f1000research.8452.1.

119. PEERE policy on data sharing on peer review. PEERE; 2017.

120. Squazzoni F, Grimaldo F, Marušić A. Publishing: Journals could share peer-review data. Nature. 2017;546:352.

121. Nature journals offer double-blind review. Nature News. 2015;518(7539):274. doi:10.1038/518274b.

122. Pulverer B. A transparent black box. The EMBO Journal. 2010;29(23):3891–3892. doi:10.1038/emboj.2010.307.

123. Ó Faoleán G. Frontiers Collaborative Peer Review: criteria to accept and reject manuscripts; 2016. Available from: https://blog.frontiersin.org/2016/09/20/the-review-process-making-decisions-on-manuscripts/.

124. Merchant S, Eckardt NA. The Plant Cell Begins Opt-in Publishing of Peer Review Reports. The Plant Cell. 2016;28(10):2343. doi:10.1105/tpc.16.00798.

125. Pourquié O, Brown K. Future developments: your thoughts and our plans. Development (Cambridge, England). 2016;143(1):1–2. doi:10.1242/dev.133355.

126. Rodgers P. Peer Review: Decisions, decisions. eLife. 2017;6:e32011. doi:10.7554/eLife.32011.