Unprofessional peer reviews disproportionately harm underrepresented groups in STEM

Nyssa J Silbiger Corresp., Equal first author, [1] , **Amber D Stubler** Equal first author, [2]

[1] Biology Department, California State University, Northridge, CA, United States

[2] Biology Department, Occidental College, Los Angeles, CA, United States

Corresponding Author: Nyssa J Silbiger
Email address: nyssa.silbiger@csun.edu

Background: Peer reviewed research is paramount to the advancement of science. Ideally, the peer review process is an unbiased, fair assessment of the scientific merit and credibility of a study; however, well-documented biases arise in all methods of peer review. Systemic biases have been shown to directly impact the outcomes of peer review, yet little is known about the downstream impacts of unprofessional reviewer comments that are shared with authors.

Methods: In an anonymous survey of international participants in STEM (science, technology, engineering, and mathematics) fields, we investigated the pervasiveness and author perceptions of long-term implications of receiving of unprofessional comments. Specifically, we assessed authors' perceptions of scientific aptitude, productivity, and career trajectory after receiving an unprofessional peer review.

Results: We show that survey respondents across four intersecting categories of gender and race/ethnicity received unprofessional peer review comments equally. However, traditionally underrepresented groups in STEM fields were most likely to perceive negative impacts on scientific aptitude, productivity, and career advancement after receiving an unprofessional peer review.

Discussion: Studies show that a negative perception of aptitude leads to lowered self-confidence, short-term disruptions in success and productivity, and delays in career advancement. Therefore, our results indicate that unprofessional reviews likely have and will continue to perpetuate the gap in STEM fields for traditionally underrepresented groups in the sciences.

Unprofessional peer reviews disproportionately harm underrepresented groups in STEM

Nyssa J Silbiger[1*†], Amber D Stubler[2†]

[1] Biology Department, California State University, Northridge, California 91330.

[2] Biology Department, Occidental College, Los Angeles, California 90041.

† This is a co-first authored paper with author order determined by ranking the most unprofessional peer reviews received by each author to date, and alphabetical order.

*Corresponding author email: nyssa.silbiger@csun.edu

Corresponding Author:
Nyssa Silbiger[1]
Biology Department, California State University, Northridge
18111 Nordhoff Street,
Northridge, CA 91330
Email address: nyssa.silbiger@csun.edu

Keywords:
Peer review; Underrepresented minorities; STEM; Intersectionality

Abstract

Background: Peer reviewed research is paramount to the advancement of science. Ideally, the peer review process is an unbiased, fair assessment of the scientific merit and credibility of a study; however, well-documented biases arise in all methods of peer review. Systemic biases have been shown to directly impact the outcomes of peer review, yet little is known about the downstream impacts of unprofessional reviewer comments that are shared with authors.
Methods: In an anonymous survey of international participants in STEM (science, technology, engineering, and mathematics) fields, we investigated the pervasiveness and author perceptions of long-term implications of receiving of unprofessional comments. Specifically, we assessed authors' perceptions of scientific aptitude, productivity, and career trajectory after receiving an unprofessional peer review.
Results: We show that survey respondents across four intersecting categories of gender and race/ethnicity received unprofessional peer review comments equally. However, traditionally underrepresented groups in STEM fields were most likely to perceive negative impacts on scientific aptitude, productivity, and career advancement after receiving an unprofessional peer review.
Discussion: Studies show that a negative perception of aptitude leads to lowered self-confidence, short-term disruptions in success and productivity, and delays in career advancement. Therefore, our results indicate that unprofessional reviews likely have and will continue to perpetuate the gap in STEM fields for traditionally underrepresented groups in the sciences.

Introduction

The peer review process is an essential step in protecting the quality and integrity of scientific publications, yet there are many issues that threaten the impartiality of peer review and undermine both the science and the scientists (Kaatz, Gutierrez, and Carnes 2014; Lee et al. 2013). A growing body of quantitative evidence shows violations of objectivity and bias in the peer review process for reasons based on author attributes (e.g. language, institutional affiliation, nationality, etc.), author identity (e.g. gender, sexuality), and reviewer perceptions of the field (e.g. territoriality within field, personal gripes with authors, scientific dogma, discontent/distrust of methodological advances) (Lee et al. 2013). The most influential demonstrations of systemic biases within the peer review system have relied on experimental manipulation of author identity or attributes (e.g. Goldberg's 1968 classic study 'Joan' vs 'John'; Goldberg 1968; Wenneras and Wold 1997) or analyses of journal-reported metrics such as number of papers submitted, acceptance rates, length of time spent in review, and reviewer scores (Fox, Burns, and Meyer 2016; Fox and Paine 2019; Helmer et al. 2017; Lerback and Hanson 2017). These studies have focused largely on the inequality of outcomes resulting from inequities in the peer review process. While these studies have been invaluable for uncovering trends and patterns, and increasing awareness of existing biases, they do not specifically assess the content of the reviews (but, see Resnik, Gutierrez-Ford, and Peddada 2008), the downstream effects that unfair, biased, and *ad hominem* comments may have on authors, and how these reviewer comments may perpetuate representation gaps in STEM fields.

In the traditional peer review process, the content, tone, and thoroughness of a manuscript review is the sole responsibility of the reviewer (the identity of whom is often protected by anonymity), yet the contextualization and distribution of reviews to authors is performed by the assigned (handling) editor at the journal to which the paper was submitted. In this tiered system, journal editors are largely considered responsible for policing reviewer comments and are colloquially referred to as the 'gatekeepers' of peer review. Both reviewers and editors are under considerable time pressures to move manuscripts through peer review, often lack compensation commensurate with time invested, experience heavy workloads, and are subject to inherent biases of their own, which may translate into irrelevant and otherwise unprofessional comments being first written, and then passed along to authors (Resnik and Elmore 2016; Resnik, Gutierrez-Ford, and Peddada 2008).

We surveyed STEM scientists that have submitted manuscripts to a peer-reviewed journal as first author to understand the impacts of receiving unprofessional peer review comments on the perception of scientific aptitude (confidence as a scientist), productivity (publications per year), and career advancement (ability to advance within the field). This study defined an unprofessional peer review comment as any statement that is unethical or irrelevant to the nature of the work; this includes comments that: 1) lack constructive criticism, 2) are directed

at the author, rather than the nature or quality of the work, 3) use personal opinions of the author/work rather than evidence-based criticism, or 4) are "mean-spirited" or cruel (for examples of comments received by survey respondents that fit these criteria see Fig. 1). The above definition was provided as a guideline for survey respondents to separate frank, constructive, and even harsh reviews from those that are blatantly inappropriate or irrelevant. Specifically, this study aimed to understand the content of the unprofessional peer reviews, the frequency at which they are received, and the subsequent impacts on the recipient's perception of their abilities. Given that psychological studies show that overly harsh criticisms can lead to diminished success (e.g., Baron 1988), we tested for the effects of unprofessional peer review on the perception of scientific aptitude, productivity, and career advancement of all respondents. Underrepresented groups in particular are vulnerable to stereotype threat, wherein negative societal stereotypes about performance abilities and aptitude are internalized and subsequently expressed (Leslie et al. 2015). Further, the combination of social categorizations may lead to amplified sources of oppression and stereotype threat (Crenshaw 1990); therefore, it is necessary to assess the impacts of unprofessional peer review comments across intersectional gender and racial/ethnic groups.

Materials & Methods

Survey Methods and Administration

The data for this study came from an anonymous survey of international members of the STEM community using Qualtrics survey software. Data were collected under institutional review board agreements and federalwide assurance at Occidental College (IRB00009103, FWA00005302) and California State University, Northridge (IRB00001788, FWA00001335). The survey was administered between February 28 and May 10, 2019 using the online-based platform in English and was open to anyone self-identifying as a member of the STEM community that published their scholarly work in a peer reviewed system. Participation in the survey was voluntary, and no compensation or incentives were offered. All respondents had to certify that they were 18 years or older and read the informed consent statement; participants were able to exit the survey at any point without consequence. Participants were recruited broadly through social media platforms, direct posting on scientific list-serves, and email invitations to colleagues, department chairs, and organizations focused on diversity and inclusive practices in STEM fields (see supplemental files for distribution methods). Targeted emails were used to increase representation of respondents. Data on response rates from specific platforms were not collected.

The survey required participants to provide basic demographic information including gender identity, level of education, career stage, country of residence, field of expertise, and racial and/or ethnic identities (see supplemental files for specific survey questions). Throughout the entire survey, all response fields included '*I prefer not to say*' as an opt-out choice. Once

demographic information was collected from participants, the study's full definition of an unprofessional peer review was presented and respondents self-identified whether they had ever received a peer review comment as first author that fit this definition. Survey respondents answering '*no*' or '*I prefer not to say*' were automatically redirected to the end of the survey, as no additional questions were necessary. Respondents answering '*yes*' were asked a series of follow-up questions designed to determine the nature of the unprofessional comments, the total number of scholarly publications to date, and the number of independent times an unprofessional review was experienced.

The perceived impact of the unprofessional reviews on the scientific aptitude, productivity, and career advancement of each respondent was assessed using the following questions: 1) To what degree did the unprofessional peer review(s) make you doubt your scientific aptitude? (1-5) 1 = not at all, 5 = I fully doubted my scientific abilities; 2) To what degree do you feel that receiving the unprofessional review(s) limited your overall productivity? Please rate this from 1-5. 1 = not at all, 5 = greatly limited number of publications per year; 3) To what degree do you feel that receiving the unprofessional review(s) delayed career advancement? Please rate this from 1-5. 1 = not at all, 5 = greatly impacted/delayed career advancement. Finally, respondents were invited to provide direct quotes of unprofessional reviews received (although the respondents were told they could remove any personal or identifying information such as field of study, pronouns, etc.). Participants choosing to share peer review quotes were able to specify whether they gave permission for the quote to be shared/distributed. Explicit permission from each respondent was received to use and distribute all quotes displayed in Fig 1. At the end of the survey, all respondents were required to certify that all the information provided was true and accurate to the best of their abilities, and that the information was shared willingly.

We recognize that there are limitations to our survey design. First, our survey was only administered in English. There also may have been non-response bias for individuals who did not experience negative comments during peer review even though any advertisement of this survey indicated that we sought input from anyone who has ever published a peer-reviewed study as first author. There could also be a temporal element, where authors who received comments more recently may have responded differently than those who received unprofessional reviews many years ago. Additionally, the order of questions was not randomized, and participants were asked to complete demographic information before answering questions about their peer review experience, which may have primed respondents to select answers that were more in line with racial or gender-specific stereotypes (Steele and Ambady 2006). In order to maintain the anonymity of our respondents, we did not ask for any bibliometric data from the authors. Given that our sample of respondents represented a diverse array of career stages, STEM fields, countries of residence, and racial/ethnic identities, we do not believe that any of the above significantly limits the interpretation of our results.

Data Analysis

We tested for the pervasiveness and downstream effects of unprofessional peer review on four intersecting gender and racial/ethnic groups: 1) women of color and non-binary people of color, 2) men of color, 3) white women and white non-binary people, and 4) white men. Due to the small number of respondents identifying as non-binary (<1% of respondents), we statistically analyzed women and non-binary genders together in a category as marganalized genders in the sciences. However, refer to supplemental Table S1 for full breakdown of responses from each gender identity so that patterns may be assessed by readers without the constraints conferred by statistical assumptions and analyses. To protect the anonymity of individuals that may be personally identified based on the combination of their race/ethnicity and gender, respondents who identified as a race or ethnicity other than white, including those who checked multiple racial and/or ethnic categories were grouped together for the statistical analysis (Fig S1). It is important to note that by grouping the respondents into four categories, the analysis captures only the broad patterns for intersectional groups and does not relay the unique experiences of each respondent, which should not be discounted.

Survey respondents (N = 1106) were given the opportunity to opt out of any question; therefore, the sample sizes were different for each statistical analysis. We tested for differences in the probability of receiving an unprofessional peer review across four intersectional groups (N = 1077) using a Bayesian logistic regression (family = Bernoulli, link = logit). Of the 642 people who indicated that they received an unprofessional peer review, 617, 620, and 618 answered the questions regarding perceived impacts to their scientific aptitude, productivity, and career advancement, respectively. We ran individual Bayesian ordinal logistic regressions (family = cumulative, link = logit) for each of the three questions to test for differences in probabilities of selecting a 1 - 5 across the four groups. All models were run using the BRMS package (Bürkner and Others 2017) in R*v3.5.2* which uses the Hamiltonian Monte Carlo algorithm in STAN (Hoffman and Gelman 2014; Team 2015). Each model was sampled from 4 chains, 4000 iterations post-warmup, and half student *t* distributions for all priors. Model convergence was assessed using Gelman-Rubin diagnostics ($\hat{R} < 1$ for all parameters; Gelman and Rubin 1992) and visual inspection of trace plots. Posterior predictive checks were visually inspected using the *pp_check()* function and all assumptions were met. Data are presented as medians and two-tailed 95% Bayesian credible intervals (BCI).

Results

We received 1106 responses from people in 46 different countries across >14 STEM disciplines (Fig. 2). Overall, 58% of all the respondents (N=642) indicated that they had received an unprofessional review, with 70% of those individuals reporting multiple instances (3.5 ± 5.8 reviews, mean ± SD, across all participants). There were no significant differences in the likelihood of receiving an unprofessional review among the intersectional groups (Fig S2); however, there were clear and consistent differences in downstream effects between groups in

perceived impacts on self-confidence, productivity, and career trajectories after receiving an unprofessional review.

White men were most likely to report no impact to their scientific aptitude (score of 1) after receiving an unprofessional peer review (P[1] =0.40 [0.34 - 0.47 95% BCI], where P[score] denotes the probability of selecting a particular score), with a 5.7 times higher probability of selecting a 1 than a 5 (fully doubted their scientific aptitude; P[5]=0.07 [0.05 - 0.09 95% BCI]). Notably, white men were 1.3, 2.0, and 1.7 times more likely to indicate no resultant doubt of their scientific aptitude than men of color (P[1]=0.30 [0.20 - 0.41 95% BCI]), white women and white non-binary people (P[1]=.20 [0.16 - 0.23 95% BCI]), and women of color and non-binary people of color (P[1]=0.23 [0.16 - 0.31 95% BCI]), respectively (Fig 3A). Together, these results indicate that receiving unprofessional peer reviews had less of an overall impact on the scientific aptitude of white men relative to the remaining three groups.

Similar patterns among intersectional groups emerged for reported impacts of unprofessional reviews on productivity (measured in number of publications per year). Specifically, women of color and non-binary people of color, white women and white non-binary people, and men of color were mostly likely to select a 3 (moderate level of perceived negative impact on productivity), whereas white men were most likely to select a 1 (no perceived impact on their productivity; Fig 3b). White men were also the least likely of all groups to indicate that receiving unprofessional reviews greatly limited their number of publications per year (P[5] = 0.06 [0.05 - 0.09 95% BCI]), which significantly differed from groups of women and non-binary people, but not men of color (Fig. 3b).

Women of color and non-binary people of color had the most distinct pattern in reported negative impacts on career advancement (Fig. 3c). Women of color and non-binary people of color had a nearly equal probability of reporting each level of impact (1-5); whereas, men of color, white women and white non-binary people, and white men had a decreasing probability of selecting scores indicative of a higher negative impact on career advancement (Fig 3c). Specifically, women of color and non-binary people of color were the most likely to select that they had significant delays in career advancement as a result of receiving an unprofessional review (P[5] = 0.20 [0.13 - 0.28 95% BCI]). Women of color and non-binary people of color were also the least likely of the groups to report no impact on career advancement as a result of unprofessional reviews (P[1] = 0.22 [0.15 - 0.31 95% BCI]).

Discussion

Our data show that unprofessional peer reviews are pervasive in STEM disciplines, regardless of race or gender, with over half of participants reporting that they had received unprofessional comments. Our study did not assess peer review outcomes of participants, but it is possible that unprofessional reviews could impact acceptance rates across groups differently because reviewer perception of competence is implicitly linked to gender regardless of content (Goldberg 1968; Kaatz, Gutierrez, and Carnes 2014; Wenneras and Wold 1997). Previous

studies have demonstrated clear differences in acceptance/rejection rates between genders (Murray et al. 2019; Symonds et al. 2006; Fox and Paine 2019) and future studies should test if receiving an unprofessional peer review leads to different acceptance outcomes depending on gender and/or race. While there were no statistical differences in the number of unprofessional reviews received among the four intersectional groups in our study, there were clear and consistent differences in the downstream impacts that the unprofessional reviews had among groups.

Overall, white men were the least likely to question their scientific aptitude, or report delays in productivity or career advancement than any other group after receiving an unprofessional review. Groups that reported the highest self-doubt after unprofessional comments also reported the highest delays in productivity. This finding corroborates studies showing destructive criticism leads to self-doubt (Baron 1988) and vulnerability to stereotype threat (Leslie et al. 2015), which has quantifiable negative impacts on productivity (Kahn and Scott 1997) and career advancement (Howe-Walsh and Turnbull 2016). Conversely, high self-confidence is related to increased persistence after failure (Baumeister et al. 2003). Therefore, scientists with a higher evaluation of their own scientific aptitude after an unprofessional review may be less likely to have reduced productivity following a negative peer review experience.

Women and non-binary people were the most likely to report significant delays in productivity after receiving unprofessional reviews. It is well known that publication rates in STEM fields differ between genders (Symonds et al. 2006; Schucan Bird 2011). Men have 40% more publications than women on average, with women taking up to 2.5 times as long to achieve the same output rate as men in some fields (Symonds et al. 2006). While our study cannot confer causality leading to diminished productivity, the results show that unprofessional reviews reinforce bias that is already being encountered by underrepresented groups on a daily basis. Other well-studied mechanisms leading to reduced productivity for women include (but are not limited to) papers by women authors spend more time in review than papers by men (Hengel 2017), men are significantly less likely to publish co-authored papers with women than with other men (Salerno et al. 2019), women receive less research funding than men in some countries (Witteman et al. 2019), and women spend more time doing service work than men at academic institutions (Guarino and Borden 2017). Women are also underrepresented in the peer review process leading to substantial biases in peer review (Goldberg 1968; Kaatz, Gutierrez, and Carnes 2014). For example, studies have shown that women are underrepresented as editors which leads to fewer refereed papers by women (Fox, Burns, and Meyer 2016; Lerback and Hanson 2017; Helmer et al. 2017; Cho et al. 2014) and that authors of all genders are less likely to recommend women reviewers (Lerback and Hanson 2017) which contributes to inequity in peer review outcomes (i.e. fewer first and last authored papers accepted by women) (Murray et al. 2019). However, some strategies, such as double-blind reviewing and open peer review, have been shown to alleviate gender inequity in publishing (Darling 2015; Budden et al. 2008; Groves 2010), although there are notable exceptions (Webb, O'Hara, and Freckleton 2008). The

difficulty in publishing and lack of productivity may contribute to the high attrition rate of women in academia (Cameron, Gray, and White 2013).

Assessment of intersectional groups has been generally overlooked in research on publication and peer review biases. Yet, traditionally underrepresented racial and ethnic groups experience substantial pressures and limitations to inclusion in STEM fields. Indeed, in our study there were significant differences, especially in perceived delays in career advancement, between white women and white non-binary people, and women of color and non-binary people of color (Fig 3c). Had we focused on only gender or racial differences, the distinct experiences of women of color and non-binary people of color would have been obscured. Because both gender and racial biases lead to diminished recruitment, retention, and attrition in the sciences (Xu 2008; Alfred, Ray, and Johnson 2019), intersectionality cannot be ignored. Our results indicate that receiving unprofessional peer reviews is an yet another barrier to equity in career trajectories for women of color and non-binary people of color, in addition to the quality of mentorship, intimidation and harassment, lack of representation, and many others (Howe-Walsh and Turnbull 2016; Zambrana et al. 2015).

Our study indicates that unprofessionalism in reviewer comments is pervasive in STEM fields. Although we found clear patterns indicating that unprofessional peer reviewer comments had a stronger negative impact on underrepresented intersectional groups in STEM, all groups had at least some members reporting the highest level of impact in every category. This unprofessional behavior often occurs under the cloak of anonymity and is being perpetuated by the scientific community upon its members. Comments like several received by participants in our study (see Fig 1) have no place in the peer review process. Interestingly, less than 3% of our participants that received an unprofessional peer review stated that the review was from an open review journal, where the peer reviews and responses from authors are published with the final manuscript (Pulverer 2010). While a recent laboratory study showed that open peer review practices led to higher cooperation and peer review accuracy (Leek et al. 2011), less is known about how transparent review practices affect professionalism in peer review comments. Our data indicate that open reviews may help curtail unprofessional comments, but more research on this topic is needed.

Individual scientists have the power and responsibility to address the occurrence of unprofessional peer reviews directly and enact immediate change. We therefore recommend the following: 1) Make peer review mentorship an active part of student and peer learning. For example, departments and scientific agencies should hold workshops on peer review ethics. 2) Follow previously published best practices in peer review (*Peer Review and Manuscript Management in Scientific Journals: Guidelines for Good Practice* 2008; e.g., Kaatz, Gutierrez, and Carnes 2014). 3) Practice self-awareness and interrogate whether comments are constructive and impartial (additionally, set aside enough time to review thoroughly, assess relevance, and re-read any comments). 4) Encourage journals that do not already have explicit guidelines for the review process to create a guide, as well as implementa process to reprimand or remove

reviewers that are acting in an unprofessional manner. For example, the journal could contact the reviewer's department chair or senior associate if they submit an unprofessional review. 5) Societies should add acceptable peer review practices to their code of conduct and a structure that reprimands or removes society members that submit unprofessional peer reviews. 6) Editors should be vigilant in preventing unprofessional reviews from reaching authors directly and follow published best practices. (D'Andrea and O'Dwyer 2017; Resnik and Elmore 2016) 7) When in doubt use the "golden rule" (review others as you wish to be reviewed).

Conclusions

Our study shows that unprofessional peer reviews are pervasive and that they disproportionately harm underrepresented groups in STEM. Specifically, underrepresented groups were most likely to report direct negative impacts on their scientific aptitude, productivity and career advancement after receiving an unprofessional peer review. While it was beyond the scope of this study, future investigations should also focus on the effect of unprofessional peer reviews on first-generation students, English as a second language, career stage, peer review in grants, and other factors that could lead to differences in downstream effects. Unprofessional peer reviews have no place in the scientific process and individual scientists have the power and responsibility to enact immediate change. However, we recognize and applaud those reviewers and editors (and there are many!) that spend a significant amount of time and effort writing thoughtful, constructive, and detailed criticisms that are integral to moving science forward.

Acknowledgements

We thank the hundreds of respondents who bravely shared some of their harshest, darkest critiques. We thank M. Siple, A. Mattheis, M. DeBiasse, J. Carroll, and the editor and two reviewers for comments and feedback on the manuscript. This is CSUN Marine Biology contribution # 296.

Data availability

The survey collected no unique personal identifying information beyond basic demographic information. However, respondent anonymity may be jeopardized through revealing specific combinations of the demographic information collected (e.g. field of study, gender, race/ethnicity). Therefore, the data that support the findings of this study are not publicly available. Data are available on request from the corresponding author (NJS); however, data will be provided in a modified format to respect the anonymity of respondents. The raw anonymized data were provided during peer review.

References

Alfred, Mary V., Sarah M. Ray, and Michele A. Johnson. 2019. "Advancing Women of Color in STEM: An Imperative for US Global Competitiveness." *Advances in Developing Human Resources* 21 (1): 114–32.

Baron, R. A. 1988. "Negative Effects of Destructive Criticism: Impact on Conflict, Self-Efficacy, and Task Performance." *The Journal of Applied Psychology* 73 (2): 199–207.

Baumeister, Roy F., Jennifer D. Campbell, Joachim I. Krueger, and Kathleen D. Vohs. 2003. "Does High Self-Esteem Cause Better Performance, Interpersonal Success, Happiness, or Healthier Lifestyles?" *Psychological Science in the Public Interest: A Journal of the American Psychological Society* 4 (1): 1–44.

Budden, Amber E., Tom Tregenza, Lonnie W. Aarssen, Julia Koricheva, Roosa Leimu, and Christopher J. Lortie. 2008. "Double-Blind Review Favours Increased Representation of Female Authors." *Trends in Ecology & Evolution* 23 (1): 4–6.

Bürkner, Paul-Christian, and Others. 2017. "Brms: An R Package for Bayesian Multilevel Models Using Stan." *Journal of Statistical Software* 80 (1): 1–28.

Cameron, Elissa Z., Meeghan E. Gray, and Angela M. White. 2013. "Is Publication Rate an Equal Opportunity Metric?" *Trends in Ecology & Evolution* 28 (1): 7–8.

Cho, Alyssa H., Shelly A. Johnson, Carrie E. Schuman, Jennifer M. Adler, Oscar Gonzalez, Sarah J. Graves, Jana R. Huebner, D. Blaine Marchant, Sami W. Rifai, Irina Skinner, Emilio M. Bruna. 2014. "Women Are Underrepresented on the Editorial Boards of Journals in Environmental Biology and Natural Resource Management." *PeerJ* 2 (August): e542.

Crenshaw, Kimberle. 1990. "Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color." *Stanford Law Review* 43: 1241.

D'Andrea, Rafael, and James P. O'Dwyer. 2017. "Can Editors Save Peer Review from Peer Reviewers?" *PloS One* 12 (10): e0186111.

Darling, E. S. 2015. "Use of Double-Blind Peer Review to Increase Author Diversity." *Conservation Biology*. https://doi.org/10.1111/cobi.12333.

Fox, Charles W., C. Sean Burns, and Jennifer A. Meyer. 2016. "Editor and Reviewer Gender Influence the Peer Review Process but Not Peer Review Outcomes at an Ecology Journal." Edited by Ken Thompson. *Functional Ecology* 30 (1): 140–53.

Fox, Charles W., and C. E. Timothy Paine. 2019. "Gender Differences in Peer Review Outcomes and Manuscript Impact at Six Journals of Ecology and Evolution." *Ecology and Evolution* 9 (6): 3599–3619.

Gelman, Andrew, and Donald B. Rubin. 1992. "Inference from Iterative Simulation Using Multiple Sequences." *Statistical Science: A Review Journal of the Institute of Mathematical Statistics* 7 (4): 457–72.

Goldberg, Philip. 1968. "Are Women Prejudiced against Women?" *Trans-Action* 5 (5): 28–30.

Groves, Trish. 2010. "Is Open Peer Review the Fairest System? Yes." *BMJ* 341 (November): c6424.

Guarino, Cassandra M., and Victor M. H. Borden. 2017. "Faculty Service Loads and Gender: Are Women Taking Care of the Academic Family?" *Research in Higher Education* 58 (6): 672–94.

Helmer, Markus, Manuel Schottdorf, Andreas Neef, and Demian Battaglia. 2017. "Gender Bias in Scholarly Peer Review." *eLife* 6 (March). https://doi.org/10.7554/eLife.21718.

Hengel, Erin. 2017. “Publishing While Female. Are Women Held to Higher Standards? Evidence from Peer Review.” https://www.repository.cam.ac.uk/handle/1810/270621.
Hoffman, Matthew D., and Andrew Gelman. 2014. “The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo.” *Journal of Machine Learning Research: JMLR* 15 (1): 1593–1623.
Howe-Walsh, Liza, and Sarah Turnbull. 2016. “Barriers to Women Leaders in Academia: Tales from Science and Technology.” *Studies in Higher Education* 41 (3): 415–28.
Kaatz, Anna, Belinda Gutierrez, and Molly Carnes. 2014. “Threats to Objectivity in Peer Review: The Case of Gender.” *Trends in Pharmacological Sciences*. https://doi.org/10.1016/j.tips.2014.06.005.
Kahn, Jeffrey H., and Norman A. Scott. 1997. “Predictors of Research Productivity and Science-Related Career Goals among Counseling Psychology Doctoral Students.” *The Counseling Psychologist* 25 (1): 38–67.
Lee, Carole J., Cassidy R. Sugimoto, Guo Zhang, and Blaise Cronin. 2013. “Bias in Peer Review.” *Journal of the American Society for Information Science and Technology* 64 (1): 2–17.
Lerback, Jory, and Brooks Hanson. 2017. “Journals Invite Too Few Women to Referee.” *Nature* 541 (7638): 455–57.
Leslie, Sarah-Jane, Andrei Cimpian, Meredith Meyer, and Edward Freeland. 2015. “Expectations of Brilliance Underlie Gender Distributions across Academic Disciplines.” *Science* 347 (6219): 262–65.
Murray, Dakota, Kyle Siler, Vincent Larivière, Wei Mun Chan, Andrew M. Collings, Jennifer Raymond, and Cassidy R. Sugimoto. 2019. “Gender and International Diversity Improves Equity in Peer Review.” *bioRxiv*. https://doi.org/10.1101/400515.
Peer Review and Manuscript Management in Scientific Journals: Guidelines for Good Practice. 2008. John Wiley & Sons.
Resnik, David B., and Susan A. Elmore. 2016. “Ensuring the Quality, Fairness, and Integrity of Journal Peer Review: A Possible Role of Editors.” *Science and Engineering Ethics* 22 (1): 169–88.
Resnik, David B., Christina Gutierrez-Ford, and Shyamal Peddada. 2008. “Perceptions of Ethical Problems with Scientific Journal Peer Review: An Exploratory Study.” *Science and Engineering Ethics* 14 (3): 305–10.
Salerno, Patricia E., Mónica Páez-Vacas, Juan M. Guayasamin, and Jennifer L. Stynoski. 2019. “Male Principal Investigators (almost) Don’t Publish with Women in Ecology and Zoology.” *PLOS ONE*. https://doi.org/10.1371/journal.pone.0218598.
Schucan Bird, Karen. 2011. “Do Women Publish Fewer Journal Articles than Men? Sex Differences in Publication Productivity in the Social Sciences.” *British Journal of Sociology of Education* 32 (6): 921–37.
Steele, Jennifer R., and Nalini Ambady. 2006. “‘Math Is Hard!’ The Effect of Gender Priming on Women’s Attitudes.” *Journal of Experimental Social Psychology* 42 (4): 428–36.
Symonds, Matthew R. E., Neil J. Gemmell, Tamsin L. Braisher, Kylie L. Gorringe, and Mark A. Elgar. 2006. “Gender Differences in Publication Output: Towards an Unbiased Metric of Research Performance.” *PLoS ONE*. https://doi.org/10.1371/journal.pone.0000127.
Team, Stan Development. 2015. “Stan: A C++ Library for Probability and Sampling, Version 2.8. 0.”
Webb, Thomas J., Bob O’Hara, and Robert P. Freckleton. 2008. “Does Double-Blind Review

473 Benefit Female Authors?" *Trends in Ecology & Evolution*.
474 Wenneras, C., and A. Wold. 1997. "Nepotism and Sexism in Peer-Review." *Nature* 387 (6631):
475 341–43.
476 Witteman, Holly O., Michael Hendricks, Sharon Straus, and Cara Tannenbaum. 2019. "Are
477 Gender Gaps due to Evaluations of the Applicant or the Science? A Natural Experiment at a
478 National Funding Agency." *The Lancet*. https://doi.org/10.1016/s0140-6736(18)32611-4.
479 Xu, Yonghong Jade. 2008. "Gender Disparity in STEM Disciplines: A Study of Faculty Attrition
480 and Turnover Intentions." *Research in Higher Education* 49 (7): 607–24.
481 Zambrana, Ruth Enid, Rashawn Ray, Michelle M. Espino, Corinne Castro, Beth Douthirt Cohen,
482 and Jennifer Eliason. 2015. "'Don't Leave Us Behind' The Importance of Mentoring for
483 Underrepresented Minority Faculty." *American Educational Research Journal* 52 (1): 40–
484 72.

485

Figure 1

Figure 1. Examples of unprofessional peer reviews from survey respondents.

Permission to publish these comments was explicitly given by respondents who certified the comments were reported accurately.

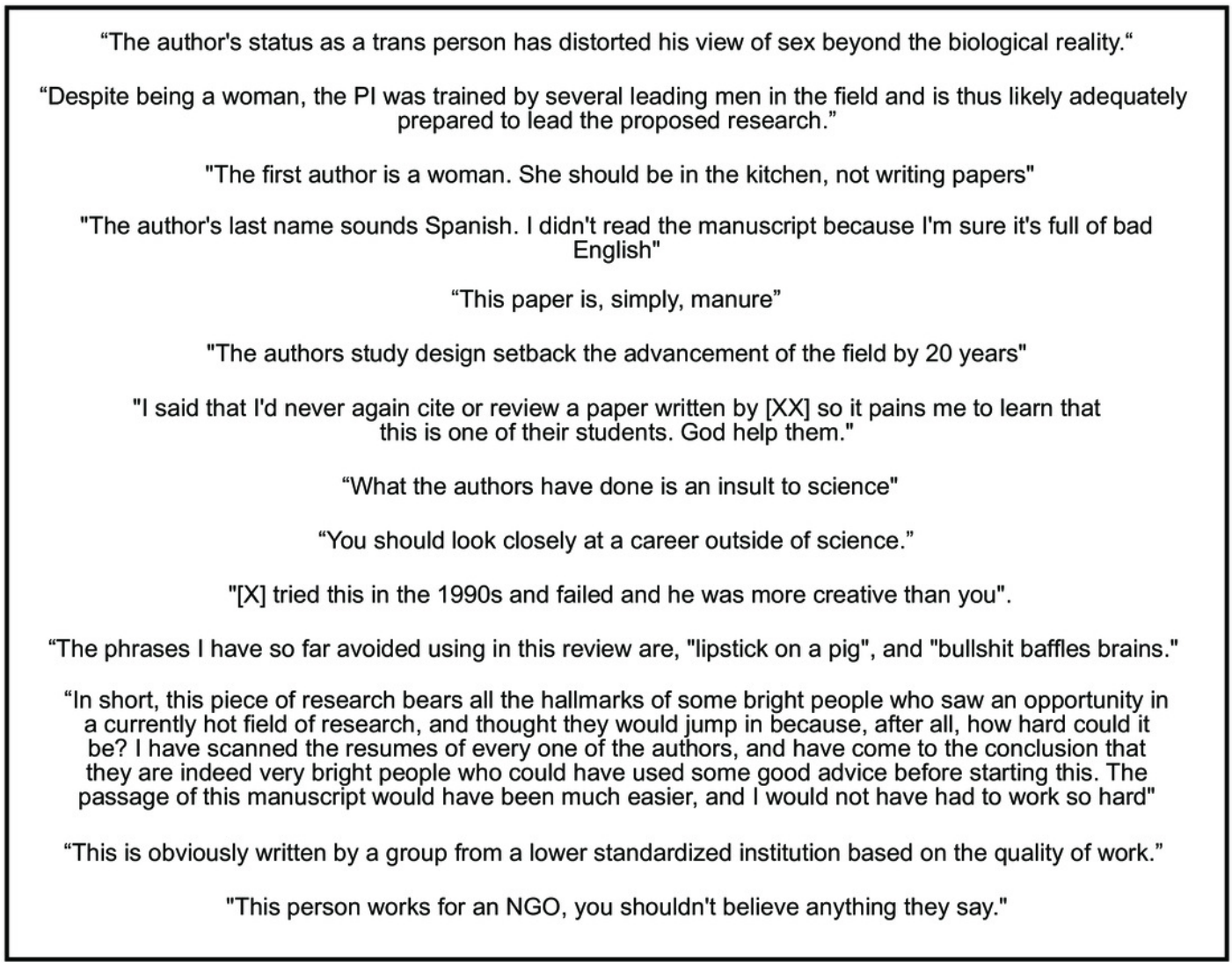


Figure 2

Figure 2: Survey demographics.

A) Representative career stages (N = 11), B) scientific disciplines (N = 14), and C) countries (N = 46) from survey participants. Color in subset (C) represents number of surveys from each country where white is 0.

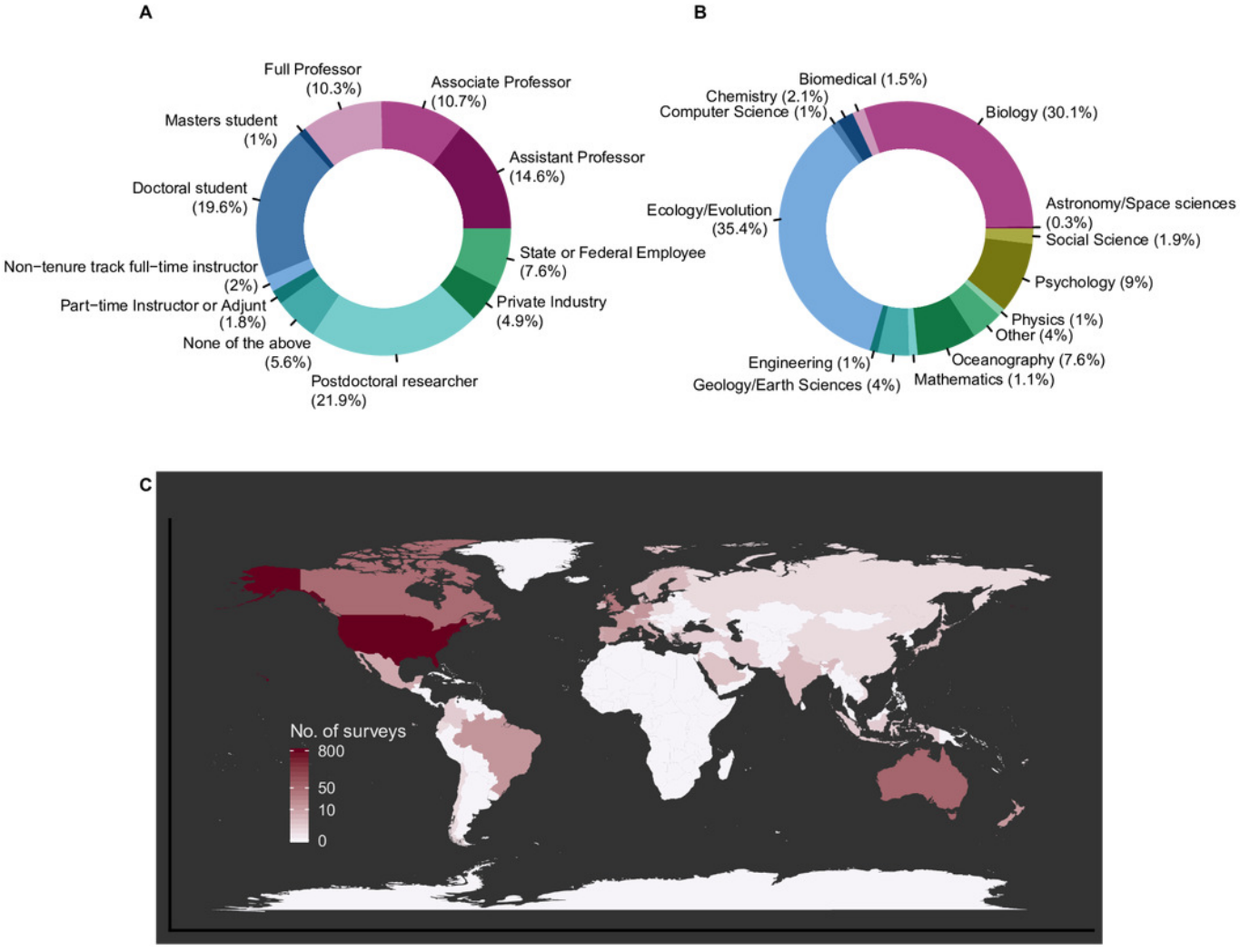
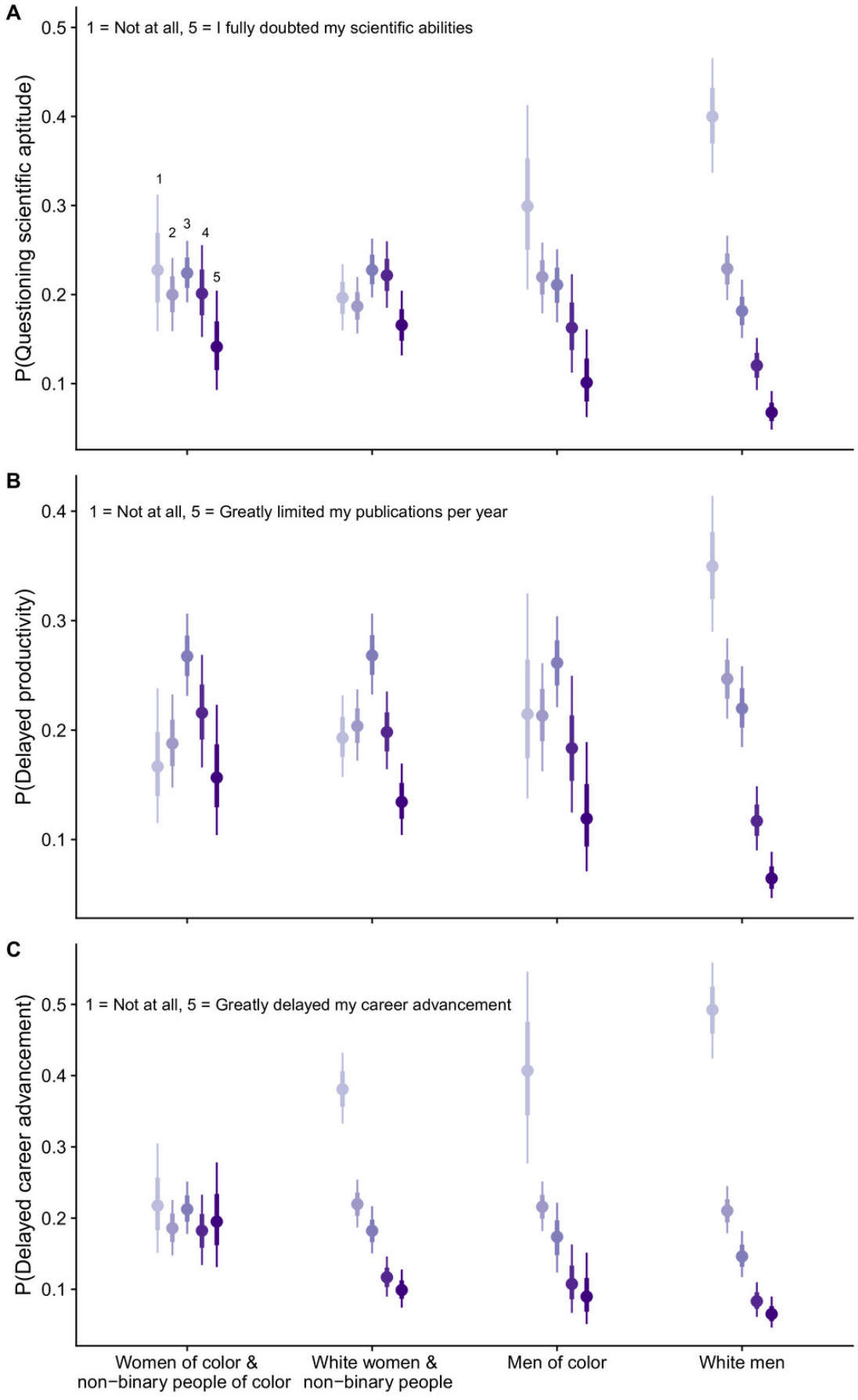


Figure 3

Figure 3: Results from Bayesian ordinal logistic regression.

Figure shows the probability of selecting a 1 - 5 for A) doubting scientific aptitude (N = 617), B) delayed productivity (N = 620), and C) delayed career advancement (N = 618) across intersectional groups after receiving an unprofessional peer review. Data are medians and two-tailed 95% Bayesian credible intervals. Colors represent level of impact with the lightest (1) as no perceived impact and the darkest (5) as the highest impact. Women and non-binary people were grouped for the statistical analysis to represent marginalized genders in STEM fields.


A
1 = Not at all, 5 = I fully doubted my scientific abilities
P(Questioning scientific aptitude)
0.5
0.4
0.3
0.2
0.1
1
2
3
4
5
B
1 = Not at all, 5 = Greatly limited my publications per year
P(Delayed productivity)
0.4
0.3
0.2
0.1
C
1 = Not at all, 5 = Greatly delayed my career advancement
P(Delayed career advancement)
0.5
0.4
0.3
0.2
0.1
Women of color & non−binary people of color
White women & non−binary people
Men of color
White men