How Gendered Is the Peer-Review Process? A Mixed-Design Analysis of Reviewer Feedback

Thomas König, *University of Mannheim, Germany*
Guido Ropers, *University of Mannheim, Germany*

ABSTRACT   A fair peer-review process is essential for the integrity of a discipline's scholarly standards. However, underrepresentation of scholarly groups casts doubt on fairness, which currently is raising concerns about a gender bias in the peer-review process of premier scholarly journals such as the *American Political Science Review (APSR)*. This study examines gender differences in *APSR* reviewing during the period 2007–2020. Our explorative analysis suggests that male reviewers privilege male authors and female reviewers privilege female authors, whereas manuscripts reviewed by both male and female reviewers indicate less gender bias. Using within-manuscript variation to address confounding effects, we then show that manuscripts reviewed by both male and female reviewers receive a more positive evaluation by female reviewers in terms of recommendation and sentiment, but they experience a marginally longer duration. Because these effects are not specific for type of authorship, we recommend that invitations to review should reflect mixed compositions of peers, which also may avoid overburdening an underrepresented group with review workload.

A fair, transparent, and nondiscriminatory peer-review process lies at the core of the scholarly publication process and builds the foundation of a trustful relationship among authors, reviewers, and editors (*American Political Science Review* 2020, v). Yet, gender publication gaps in leading political science journals—including the *American Political Science Review (APSR)*—cast doubt on the integrity of this process (Breuning et al. 2018; Brown and Samuels 2018; Teele and Thelen 2017). This aligns with recurrently voiced concerns across disciplines questioning whether editors and reviewers from different scholarly groups (e.g., qualitative versus quantitative, senior versus junior, and male versus female) hold manuscripts of authors from other groups to different standards (Card et al. 2020; Lee et al. 2013).

Most studies emphasize the role of editors as the driver of group-specific biases in editorial processes (Squazzoni et al. 2020; Teele and Thelen 2017). Others correlate submission with publication rates and point to systematically lower submission rates of female authors (Brown and Samuels 2018; König and Ropers 2018). Because lower submission rates might be caused by a higher risk aversion of women (Djupe, Smith, and Sokhey 2019) and/or perceptions that the peer-review process is organized at the expense of female work (Brown et al. 2020), explorative or correlational insights into gender bias risk omitted variable bias from unobserved confounders. In our mixed-design analysis of the *APSR*, we first explore reviewer bias across authorship type and then use within-manuscript variation from manuscripts that were evaluated by both male and female reviewers to identify the gender-specific impact of reviewers who provide feedback to both editors and authors. Stated differently, if men and women assess the same manuscript to other standards, a nonrandom assignment of reviewers to manuscripts based on gender may set the foundation for gender bias in the review process.

For the period 2007–2020, our explorative findings suggest that a gender bias exists: reviews for manuscripts by male authors reviewed only by men are, on average, more positive than for other

Thomas König [iD] *is professor of political science at the University of Mannheim. He can be reached at tkoenig@uni-mannheim.de.*

Guido Ropers [iD] *is a doctoral researcher at the University of Mannheim. He can be reached at gropers@mail.uni-mannheim.de.*

types of authors, whereas reviews are more positive for female-authored manuscripts than for other types of authors if only women are reviewing. To enhance the internal validity of our findings—and compared to studies that do not consider unobserved confounders in the peer-review process—we then control for the quality of manuscripts to examine gender-specific reviewer

standards conditional on their gender. It is a disadvantage that women are facing in many areas of the academic profession (e.g., Hengel 2017; Mengel, Sauermann, and Zölitz 2019; Sarsons 2017).

Our empirical analysis of the *APSR*'s double-blind peer-review process encompasses submission and review data for the period 2007–2020. We restrict our sample to manuscripts submitted after

A fair, transparent, and nondiscriminatory peer-review process lies at the core of the scholarly publication process and builds the foundation of a trustful relationship among authors, reviewers, and editors.

differences in feedback for the same manuscript. Our analysis considers not only reviewer recommendations but also other feedback such as length, sentiment, and duration of reviews. We find that women, who constitute a smaller share of the reviewer pool, provide more positive feedback than men independent of authorship. In addition to a 7% higher likelihood for a non-reject recommendation, the review tone of female reviewers is more positive. However, their reviews take a few days longer but are independent of authors' gender.

Given the feedback differences between men and women, we recommend ensuring gender diversity in the peer-review process by having at least one woman and one man as manuscript reviewers. This recommendation attempts to consider direct and indirect effects of gender bias by taking into consideration a gender-specific distribution of available reviewers. If the pool of reviewers consists of potential authors, it is likely that—due to their lower share—female authors will have less time for conducting, submitting, and publishing their own research when they spend much of their time to review manuscripts. Inviting more than one female reviewer can increase slightly the chances of a manuscript surviving the peer-review process. However, without compensation, there is a risk to increase the gender gap in top-ranked science journals such as the *APSR*, which already can be inferred from gender-specific rejections of review invitations.

REVIEWER RECOMMENDATIONS AND GENDER BIAS

A double-blind review process as implemented by the *APSR*, in which neither the reviewers nor the authors are aware of the others' group-specific identity, ideally prevents biased reviewing

July 1, 2007, and for which a first decision was made before May 31, 2020 (König and Ropers 2021). To obtain information on the gender of authors and reviewers, we coded first names using the genderizeR-package (Wais 2016).[1] Regarding authorship, we differentiate among women only, men only, and mixed-gender submissions of all authors, whereby we aggregate solo and co-authorship by gender.

Table 1 lists descriptive statistics by gender. In the overall sample, 65% of submissions were authored by solo male authors or all-male teams; 16% of manuscripts were submitted by solo female authors or all-female teams; and 19% of submissions were authored by mixed-gender teams. Among invited reviewers, women comprise 29% of all reviewers, which corresponds to their 28% share of tenured political science faculty (Alter et al. 2020). The gender distribution among "submitting" reviewers—that is, reviewers who not only accepted a review request but also submitted their report—was slightly lower with 27% of reviewers being women.

Recommendations of reviewers (i.e., accept, minor revision, major revision, or reject) are important signals for the decision of editors to continue with peer review. Assuming that we can recode them on a standard ordinal scale (with equiprobable occurrence probability), we would determine, for example, that an "accept" recommendation is clearly better than a "minor-revision" recommendation. However, this comparison becomes more complicated when the number of reviews increases. To make the set of recommendations comparable across manuscripts with different numbers of reviews, we followed Bravo et al. (2018) and calculated a review score for each manuscript. This measures the "value" of a

Given the feedback differences between men and women, we recommend ensuring gender diversity in the peer-review process by having at least one woman and one man as manuscript reviewers.

behavior based on author characteristics. Several studies find that double-blindness increases submissions from women (Budden et al. 2008) and provides a fairer review process compared to a single-blind review process with reviewer knowledge about the author (Tomkins, Zhang, and Heavlin 2017). However, the level of anonymity of the double-blind review process can be questioned in practice as reviewers may be able, for example, to infer the identity of authors from "gendered research agendas" (Key and Sumner 2019). As a result, authors may be held to different

given combination of recommendations by counting all other possible combinations for a given number of recommendations that are both clearly better and clearly worse than those that a manuscript received. For example, for a manuscript with an {accept, major revision} recommendation set, there are two recommendations that clearly are better: {accept, accept} and {accept, minor revision}; six combinations that clearly are worse; and one combination, {minor revision, minor revision}, that is unclear.[2] Being bound between 0 and 1, the review score allows a

Descriptive Statistics: Gender in the *APSR*

		N	Share
Authorship	Mixed-Gender Team	2,258	19%
	Male Author(s)	7,669	65%
	Female Author(s)	1,867	16%
Invited Reviewers	Male	32,684	71%
	Female	13,184	29%
Submitting Reviewers	Male	17,418	73%
	Female	6,592	27%

Note: Sample includes manuscripts submitted after July 1, 2007, that received a first decision before May 31, 2020.

comparison of scores for manuscripts that received a different number of review recommendations and is calculated as follows:

$$\text{review score} = \frac{\#worse}{\#better + \#worse}$$

Figure 1 displays the average review scores for the three gender compositions of authors from three different gender-reviewer compositions: all male, mixed-gender, and all female. Two seemingly opposing patterns stand out. First, the average score depicted on the left and right categories in the graph show that if the reviewer composition consists of only the same gender, there is a strong indication of gender bias. That is, when a manuscript is reviewed only by men, male authors receive higher scores compared to scores of manuscripts written by mixed-gender and female authors. Conversely, higher review scores for female authors are associated with female reviewers compared to the scores of male and mixed-gender authors. Both trends point toward the existence of gender bias in the peer-review process. They suggest that both male and female authors are held to different standards conditional on whether they are being reviewed by only male or only female reviewers.

Second, figure 1 also shows that the category of mixed reviewers in the middle of the graph is a less gender-biased pattern. There is little variation in the review scores conditional on the type of authors (i.e., 2,673 male, 968 mixed-gender, and 796 female authors) if manuscripts were reviewed by at least one male and one female reviewer. This is an important finding because it suggests that manuscripts—independent of the type of authorship—are held at similar standards when they were reviewed by mixed-gender reviewers and if we assume that the manuscript quality is comparable. Methodologically, however, we cannot rule out that the differences are driven at least partially by selection effects (Helmer et al. 2017). In the *APSR* case, manuscripts written by women have a much larger percentage of female reviewers (36%) compared to both mixed-gender teams (30%) and male authors (25%). Another selection effect concerns field-specific particularities, such as "methodological proclivities" (Teele and Thelen 2017, 433). All of this suggests that we need to control not only for the type of author (i.e., solo/team, male/female, or mixed gender) but also for the "intrinsic quality of manuscripts" (Squazzoni et al. 2020, 4) to ensure the internal validity of findings.

Average Review Scores by Different Reviewer Gender Compositions Conditional on Authors' Gender Composition
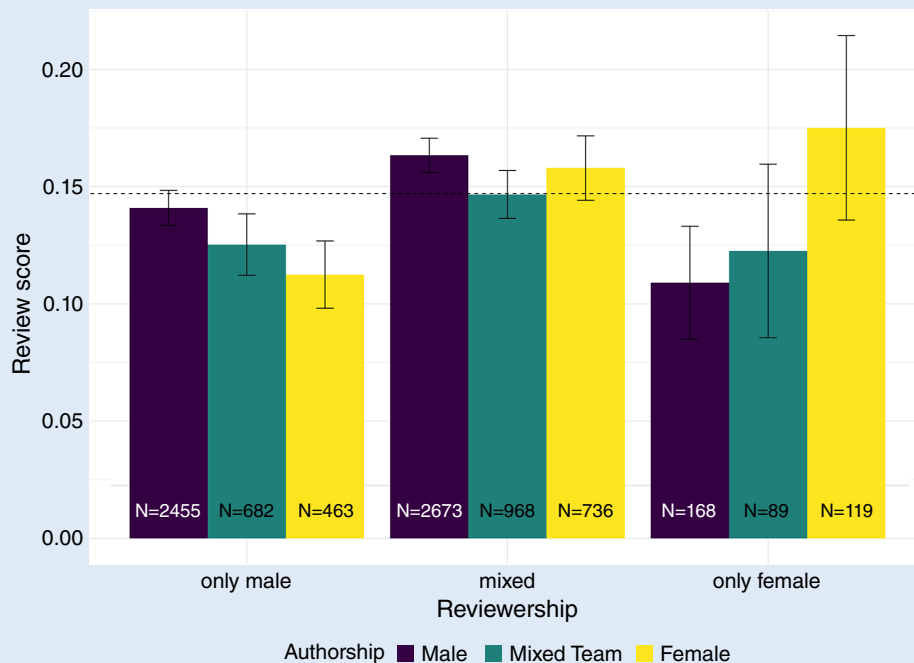
Note: The dotted horizontal line indicates the average review score in the sample.

CONTROLLING FOR MANUSCRIPT QUALITY AND REVIEWER FEEDBACK

To control for the quality and other invariant manuscript-level confounders, we turn to a fixed-effects approach at the manuscript level that focuses on the category of mixed-gender reviewed

variables holding constant quality and other invariant characteristics at the manuscript level. In each model specification, the observations are weighted by the inverse number of reviews for the manuscript and standard errors are clustered at the manuscript level. The baseline is the feedback from male reviewers.

> *When a manuscript is reviewed only by men, male authors receive higher scores compared to scores of manuscripts written by mixed-gender and female authors.*

manuscripts. This category constitutes a subsample in which each manuscript received reviews from at least one female and one male reviewer. The resulting regression coefficient for female reviewer estimates the average difference between female and male reviewers using variation from the same manuscript.

We focused on four observable measures of reviewer feedback. Our first measure is the reviewers' recommendation to either reject or proceed with the peer-review process—often considered as the key signal of reviewers to editors. We used non-reject recommendations as the main measurement for the level of reviewer support. A dummy variable was coded as 1 if a reviewer provided a non-reject recommendation (i.e., major revision, minor revision, or accept).

Recommendations are only one type of reviewer feedback. The second measure and one of the main benefits that authors can take from the review process—which ends with a rejection for more than 90% of submitted manuscripts in premier scholarly outlets such as the *APSR*—is the substantive feedback of the reviewers who are helping them to improve their research (*APSR* 2020). Although it is difficult to measure the overall substance of a review, one measure—admittedly very crude—is its length (i.e., number of words).[3] This may serve as a proxy for how seriously a reviewer assesses a manuscript and how much professional feedback is (quantitatively) provided to improve an author's research.

The third measure examines the sentiment or tone of the submitted reviews. Specifically, if reviews that women receive are systematically more negative coming from the larger share of the male-reviewer pool, it may discourage them to continue their research and resubmit to a premier scholarly outlet. We used a dictionary approach and calculated a sentiment score defined as the difference between the percentage of words with positive and negative connotations. Higher values measured a more favorable tone. The classification into positively and negatively associated words is based on the Non-Commercial Research Use Word–Emotion Association Lexicon (Mohammad and Turney 2013) in the quanteda package (Benoit et al. 2018).

The fourth relevant measure of reviewer feedback, in particular for young professionals, is the time it takes for reviewers to submit their reviews. Hengel (2017) found that the editorial decision-making process is longer for female authors than for male authors. In contrast, Card et al. (2020) assessed reviewer differences in the time it takes for reviewers to submit their reports and did not find any differences. To examine whether this duration differs for male and female authors at the *APSR*, we focused on the time that a reviewer accepted a review until he or she submitted the review.

Table 2 presents the average difference between female and male reviewers for each of the four dependent reviewer-feedback

We first examine the probability of receiving a non-reject recommendation using a linear probability model. According to the estimates shown in table 2, column 1, female reviewers are about 3 percentage points more likely to give a positive (i.e., non-reject) recommendation on average (standard error: 1 percentage point). Given an average probability for a non-reject recommendation of about 42%, the estimated average effect constitutes an increase of about 7%.

Unlike for recommendations, however, we did not observe a statistical difference in review length between male and female reviewers, on average. Moreover, the substantive size of the point estimate of about 17 fewer words (standard error: 10) also is small given an average text length of about 880 words for manuscripts with at least one male and one female reviewer.

Not only the length of a review but also its tone may affect how authors perceive their submission experience. In this regard, female reviewers seem to be more positive in their tone when writing a review, as shown in the estimates in table 2, column 3. Relative to an average sentiment score of about 3.2, the point estimate of 0.16 corresponds to an increase of about 5%.[4]

Regarding review duration, small differences between female and male reviewers become apparent. Female reviewers take about 1.6 days longer, on average, to submit their review (standard error: 0.4). However, this average difference of less than two days is negligible given an average review duration of about 33 days for the *APSR*.

In summary, our analyses of reviewer feedback indicate gender differences among reviewers, on average, for three of the four examined outcome measures. We find a higher share of non-reject recommendations and a more positive review tone by women. Moreover, women seem to take about one to two days longer for their review, independent of authorship type, but we do not find differences in review text length. The next section examines whether these differences are amplified conditional on authorship type.

THE ROLE OF AUTHORSHIP TYPE

In the final step of our analysis, we examined whether men and women hold manuscripts to different standards conditional on authorship type. To do so, we interacted reviewer gender with authorship type in our fixed-effects specification. Table 3 presents the coefficient estimates. The interaction effects capture aggregate differences in the assessment between female and male reviewers for different authorship types relative to the average difference between men and women that is estimated for the reference group (i.e., mixed-gender-team submissions). For a causal interpretation, we assumed that the differences between men and women do not change to a different degree for other reasons.

Table 2
OLS Regression Results: Gender-Reviewing Feedback

	Non-Reject Recommendation	Review Length	Sentiment	Duration
Female Reviewer	0.03***	−17.46	0.16***	1.60***
	(0.01)	(10.02)	(0.03)	(0.37)
Manuscript Fixed Effects	Yes	Yes	Yes	Yes
Number of Observations	23,948	20,267	20,267	23948
Number of Manuscripts	8,294	7,069	7,069	8,294

Notes: *** p<0.001, ** p<0.01, * p<0.05. Standard errors are clustered at the manuscript level.

Table 3
OLS Regression Results: Gender-Reviewing Feedback and Authorship Type

	Non-Reject Recommendation	Review Length	Sentiment	Duration
Female Reviewer	0.04*	11.86	0.15*	1.80*
	(0.02)	(21.45)	(0.06)	(0.78)
Male Authors * Female Reviewer	−0.01	−35.57	0.02	0.03
	(0.02)	(−25.02)	(0.07)	(0.91)
Female Authors * Female Reviewer	0.01	−44.09	0.01	1.06
	(0.03)	(32.08)	(0.09)	(1.19)
Manuscript Fixed Effects	Yes	Yes	Yes	Yes
Number of Observations	23,948	20,267	20,267	23,948
Number of Manuscripts	8,294	7,069	7,069	8,294

Notes: ***p<0.001, ** p<0.01, * p<0.05. Reference groups are male reviewers and mixed-gender-team authorship. Standard errors are clustered at the manuscript level.

The estimates in the reference group—mixed-gender-team authorship type—correspond to the average estimates shown in table 2. Accordingly, women are more likely than men to provide a non-reject recommendation, are more positive in their writing, and take slightly longer to submit a review. Moreover, for none of the four outcome measures do we find statistically significant interaction effects for other authorship types—that is, statistical differences in the estimate for a female reviewer relative to the reference group (i.e., mixed-gender-team authorship type).

Figure 2 illustrates the corresponding average marginal effects of a female reviewer conditional on authorship type. Two patterns are worth mentioning despite the lack of statistically significant interaction effects. First, the point estimate for the probability of receiving a non-reject recommendation from a female reviewer versus a male reviewer is larger for female-authored manuscripts than for male-authored and mixed-gender-team-authored manuscripts. Second, although female reviewers take about two days longer than men to review both male and mixed-gender-team-authored manuscripts, the marginal effect is statistically indistinguishable from zero for the subgroup of female-authored manuscripts.[5]

Nevertheless, the findings do not allow us to conclude that the type of feedback from men or women depends on authorship type (i.e., gender composition).[6] Stated differently, the findings suggest that male and female authors are not held to different standards by men and women when they are reviewed by a mixed-gender composition of reviewers.

DISCUSSION AND CONCLUSION

Although a fair peer-review process is essential for a trustful relationship among editors, reviewers, and authors, the under-representation of women in many premier scholarly journals casts doubt on the integrity of this process and raises concerns about gender bias. Our analysis of reviewer feedback at the *APSR* detects a biased pattern in the recommendations of reviewers for authors of the same gender. Using a fixed-effects approach to control for selection effects in reviewer assignment and unobserved confounders at the manuscript level, we also find differences in reviewer feedback between men and women who review the same manuscript. On average, women are more likely to provide a non-reject recommendation and also are more positive than men in their review tone. Moreover, women seem to take about one to two days longer for their review, independent of authorship type, although we find little differences in review text length. However, our results do not suggest that any of these differences in reviewer feedback are conditional on the authorship type for the subsample of manuscripts with mixed-gender reviewer composition.

This study makes two main contributions to the literature on gender bias in premier scholarly journals. First, it adds to an earlier study examining various editorial outcomes in the *APSR*, including reviewer recommendations. König and Ropers (2018) documented that female reviewers are more likely to provide non-reject recommendations than male reviewers in line with the findings of this article. However, their analysis additionally

Figure 2

Average Marginal Effect of a Female Reviewer Relative to a Male Reviewer Conditional on Authorship Type
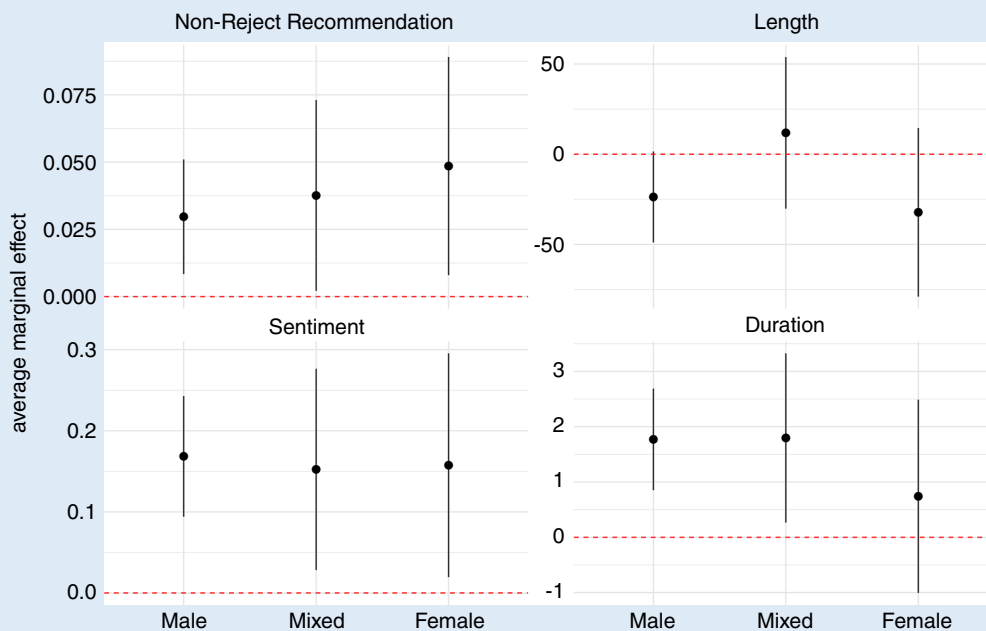
provided suggestive evidence that women might be particularly friendly toward all-female-author teams (König and Ropers 2018, 850). Yet, these authors acknowledge that the across-manuscript comparison of their research design "lack[s] objective information on the quality of manuscripts" (König and Ropers 2018, 851) to confidently exclude bias from unobserved confounders that may be driving the reviewer differences between men and women. This caveat is addressed through the fixed-effects approach in this study that focuses on within-manuscript variation to hold the quality of manuscripts constant.

Second, our findings align with recent causal analyses in other disciplines that also control for manuscript-level confounders in reviewer recommendations (Card et al. 2020). By looking closer at different types of reviewer feedback, our study highlights that it is possible to reduce gender bias in several outcomes when reviewers have a mixed-gender background. In this category, we do not find an association between the gender of authors and the overall feedback they receive. If we assume that no confounders exist in the assignment of manuscripts, then our findings suggest that bias can increase or decrease the publication chances of specific groups of authors.

Given the general positivity of female reviewers and that the same-gender reviewer bias for male-authored submissions is slightly smaller than the relative difference in review scores that female-authored submissions receive from all-female compared to all-male reviewer compositions, an exclusive assignment of female reviewers to manuscripts of female authors may reduce the gender publication gap.[7] However, if such an assignment comes at the expense of increased professional services for women in the discipline, this risks disproportionately keeping women from conducting and submitting their own research to

premier scholarly outlets. We have already observed a decrease in women who accept invitations to review for the *APSR*. Although the acceptance rate of men and women was almost the same until 2011, there is an incrementally increasing gap emerging with a 9-percentage-points lower acceptance of invited female reviewers for manuscripts submitted in 2019 (see also Breuning et al. 2015).

In comparison, our manuscript fixed-effects analysis holding constant confounders like manuscript quality suggests that the assignment of a mixed-gender composition with at least one male and one female reviewer provides a fair peer-review process for all types of authors. Given the high predictive power of the review score for a manuscript's editorial outcome (as shown in table S4, model 2, in the online appendix), and considering a fair review process for all submitting authors, it is possible to aim for an even distribution of women (or other underrepresented scholarly groups) across manuscripts that are sent out for review. For example, with approximately 30% women in political science (Alter et al. 2020) and typically three reviewers per manuscript in premier journals, this implies inviting at least one woman as reviewer, on average, without the risk of overburdening female scholars.

DATA AVAILABILITY STATEMENT

Research documentation and data that support the findings of this study are openly available on the *PS: Political Science & Politics* Harvard Dataverse at https://doi.org/10.7910/DVN/LMDUEQ.

SUPPLEMENTARY MATERIALS

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/S1049096521000937. ∎

NOTES

1. We hand-coded cases for which gender was not identified by the algorithm or for which the gender of the first and middle names contradicted each other. (Accessed on December 20, 2018.)

2. See table S1 in the online appendix for the full "potential recommendation set" (Bravo et al. 2018, 103) for two recommendations.

3. We could count only the number of words of reviews that were submitted via our online submission system. We excluded manuscripts from our analysis for which at least one written review was missing.

4. We also separately examined the percentage of positive and negative words (see table S2 in the online appendix). The results suggest that the difference in the sentiment score is driven by the usage of positive words. In addition, we did not find differences in the usage of words associated with disgust—indicated by dismissive comments—between men and women neither on average nor conditional on authorship type.

5. However, this may be a result of the relatively small number of female-authored manuscripts in the sample.

6. Online appendix table S3 presents the results from our main analysis differentiating between team and solo submissions. The results mirror those presented in tables 2 and 3 in both substantive and statistical terms.

7. This aligns with the manuscript-level results; every additional female reviewer further increases the average review score, as shown in figure S1 and table S4, model 1, in the online appendix.

REFERENCES

Alter, Karen J., Jean Clipperton, Emily Schraudenbach, and Laura Rozier. 2020. "Gender and Status in American Political Science: Who Determines Whether a Scholar Is Noteworthy?" *Perspectives on Politics* 18 (4): 1–20.

American Political Science Review (*APSR*). 2020. "Notes from the Editors." *American Political Science Review* 114 (2): v–viii.

Benoit, Kenneth, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Müller, and Akitaka Matsuo. 2018. "quanteda: An R Package for the Quantitative Analysis of Textual Data." *Journal of Open Source Software* 3 (30): 774.

Bravo, Giangiacomo, Mike Farjam, Francisco Grimaldo Moreno, Aliaksandr Birukou, and Flaminio Squazzoni. 2018. "Hidden Connections: Network Effects on Editorial Decisions in Four Computer Science Journals." *Journal of Informetrics* 12 (1): 101–12.

Breuning, Marijke, Jeremy Backstrom, Jeremy Brannon, Benjamin Isaak Gross, and Michael Widmeier. 2015. "Reviewer Fatigue? Why Scholars Decline to Review Their Peers' Work." *PS: Political Science & Politics* 48 (4): 595–600.

Breuning, Marijke, Benjamin Isaak Gross, Ayal Feinberg, Melissa Martinez, Ramesh Sharma, and John Ishiyama. 2018. "Clearing the Pipeline? Gender and the Review Process at the *American Political Science Review*." *PS: Political Science & Politics* 51 (3): 629–34.

Brown, Nadia E., Yusaku Horiuchi, Mala Htun, and David Samuels. 2020. "Gender Gaps in Perceptions of Political Science Journals." *PS: Political Science & Politics* 53 (1): 114–21.

Brown, Nadia E., and David Samuels. 2018. "Introduction to Gender in the Journals, Continued: Evidence from Five Political Science Journals." *PS: Political Science & Politics* 51 (4): 847–48.

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

Card, David, Stefano DellaVigna, Patricia Funk, and Nagore Iriberri. 2020. "Are Referees and Editors in Economics Gender Neutral?" *Quarterly Journal of Economics* 135 (1): 269–327.

Djupe, Paul A., Amy Erica Smith, and Anand Edward Sokhey. 2019. "Explaining Gender in the Journals: How Submission Practices Affect Publication Patterns in Political Science." *PS: Political Science & Politics* 52 (1): 71–77.

Helmer, Markus, Manuel Schottdorf, Andreas Neef, and Demian Battaglia. 2017. "Gender Bias in Scholarly Peer Review." Edited by Peter Rodgers. *eLife* 6:e21718.

Hengel, Erin. 2017. *Publishing While Female. Are Women Held to Higher Standards? Evidence from Peer Review.* Working Paper. Accessed February 5, 2018. www.repository.cam.ac.uk/handle/1810/270621.

Key, Ellen M., and Jane Lawrence Sumner. 2019. "You Research Like a Girl: Gendered Research Agendas and Their Implications." *PS: Political Science & Politics* 52 (4): 663–68.

König, Thomas, and Guido Ropers. 2018. "Gender and Editorial Outcomes at the *American Political Science Review*." *PS: Political Science & Politics* 51 (4): 849–53.

König, Thomas, and Guido Ropers. 2021. "Replication Data for: How Gendered Is the Peer-Review Process? A Mixed-Design Analysis of Reviewer Feedback." https://doi.org/10.7910/DVN/LMDUEQ.

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.

Mengel, Friederike, Jan Sauermann, and Ulf Zölitz. 2019. "Gender Bias in Teaching Evaluations." *Journal of the European Economic Association* 17 (2): 535–66.

Mohammad, Saif M., and Peter D. Turney. 2013. "Crowdsourcing a Word-Emotion Association Lexicon." *Computational Intelligence* 29 (3): 436–65.

Sarsons, Heather. 2017. "Recognition for Group Work: Gender Differences in Academia." *American Economic Review* 107 (5): 141–45.

Squazzoni, Flaminio, Giangiacomo Bravo, Pierpaolo Dondio, Mike Farjam, Ana Marusic, Bahar Mehmani, Michael Willis, Aliaksandr Birukou, and Francisco Grimaldo. 2020. *No Evidence of Any Systematic Bias Against Manuscripts by Women in the Peer-Review Process of 145 Scholarly Journals.* SocArXiv, February 28. Accessed June 21, 2021. https://osf.io/preprints/socarxiv/gh4rv.

Teele, Dawn Langan, and Kathleen Thelen. 2017. "Gender in the Journals: Publication Patterns in Political Science." *PS: Political Science & Politics* 50 (2): 433–47.

Tomkins, Andrew, Min Zhang, and William D. Heavlin. 2017. "Reviewer Bias in Single- Versus Double-Blind Peer Review." *Proceedings of the National Academy of Sciences* 114 (48): 12708–13.

Wais, Kamil. 2016. *genderizeR: Gender Prediction Based on First Names.* https://CRAN.R-project.org/package=genderizeR.