

The Review Process and the Citation Gap: The Role of the Editor's Nudge

Chris W. Bonneau, University of Pittsburgh
Kristin Kanthak, University of Pittsburgh
Amanda Leifson, University of Pittsburgh
Shane M. Redman, University of Pittsburgh

Political scientists have long been interested in how “gendered institutions”—institutions for which gender at least partially affects outcomes—operate in the political arena. Yet we know relatively little about how our discipline’s own institutions might be gendered. In the current project, we examine how gender operates within the institution of the journal review process. We implement a unique double-blind experiment embedded in the review process at a real academic journal in which the editor encourages a randomly selected subset of reviewers to provide additional citations, including citations to their own work. The nudge did not increase the probability that a reviewer would cite a woman, but it did increase the number of women cited given that the reviewer cited at least one woman, an expected increase of 0.05 citations to women per review. Furthermore, the nudge actually exacerbated the self-citation gender gap.

Gendered institutions” are those in which women are excluded or segregated, often through processes that are not explicitly about gender (Acker 1992). And while political scientists have given a great deal of attention to how gendered institutions such as legislatures affect political outcomes (e.g., Kerevel and Atkeson 2013), we know relatively little about how gendered institutions operate in our own discipline.

Despite this, we do know that women face significant barriers to advancement in the academy, and part of the reason why women fall behind men in academia is the so-called citation gender gap (Mitchell, Lange, and Brus 2013; Ostby et al. 2013). Citations tend to be the “coin of the realm” in the academy (Symonds et al. 2006), but papers with women as first author receive systematically fewer citations across myriad disciplines (Larivière et al. 2013), including political science. Indeed, Maliniak, Powers, and Walter (2013) find that male-authored international relations articles received about 4.8 more citations than female-authored articles, a gap that

does not go away once a host of explanatory variables are included. These patterns can have long-term effects on women’s careers (Alper 1993; Symonds et al. 2006). Furthermore, simply increasing the number of women in the discipline is unlikely on its own to prompt an improvement in how women are treated (Kanthak and Krause 2012). In the journal we formerly edited (Bonneau and Kanthak as coeditors, Leifson and Redman as editorial assistants), *State Politics and Policy Quarterly*, in all 300 published articles from 2001 to 2014, the average number of citations was 48.0, but only 11.9 of these had a female author. The average article cited 77.9 male scholars but only 13.6 female scholars.

There are several explanations for this citation gender gap. Women do not get the same benefit from collaborating that men do (Teele and Thelen 2017); women are less likely to appear on syllabi, particularly those prepared by men (Colgan 2017); and women are much less likely than men to cite their own previous work (King et al. 2017). Moreover, the practice of self-citing increases citations from others, with no penalty,

Chris W. Bonneau is professor of political science, Kristin Kanthak is associate professor, and Amanda Leifson and Shane M. Redman are PhD candidates at the University of Pittsburgh, Pittsburgh, PA 15217. Contact the corresponding author, Kristin Kanthak, at kanthak@pitt.edu.

The study reported in this article was conducted in compliance with relevant laws and was deemed exempt by the University of Pittsburgh institutional review board. Data and supporting materials necessary to reproduce the numerical results in the article are available in the *JOP* Dataverse (<https://dataverse.harvard.edu/dataverse/jop>). An online appendix with supplementary material is available at <https://doi.org/10.1086/710326>.

for even the most habitual self-citers (Fowler and Aksnes 2007). Because we know that gendered institutions are often characterized by the “appearance of gender neutrality” (Acker 1992, 568), the review process may be gendered if it encourages these behaviors, even if it does not specifically mention gender. We investigate the review process as a gendered institution, and explore ways to mitigate its effects.

THE EDITOR’S NUDGE

The current article assesses the role of a journal editor’s “nudge,” an encouragement to reviewers to increase citations, including citations to oneself. Between 2015 and 2017, we inserted an experiment into the review process of a peer-reviewed political science subfield journal, *State Politics and Policy Quarterly*. The journal has an impact factor of 1.675 and receives between 100 and 120 submissions per year. The acceptance rate hovers around 20%, and a bit over half of all reviewers recommend rejection. Most rejected papers are subsequently published in other outlets. We provide a detailed account of the review process in the next section.

The “nudge” consisted of a message in the initial invitation to review asking reviewers to provide citations to any relevant work they feel is missing in the manuscript, including references to their own work. We specifically do not mention gender in the nudge because we are interested in how putatively “gender neutral” gendered institutions operate. Rather, the nudge is meant to prompt reviewers to slow down their thinking in the sense of Kahneman (2011), to go beyond the easiest-to-recall (and perhaps more male) exemplars to think of other work that might be more likely to be authored by women. Similarly, drawing on the long literature on women in politics, we know that women are less likely to “tout their own qualifications” (Lawless and Fox 2005, 97) and more likely to need encouragement to run for office, implying that women may also benefit from a similar invitation to self-cite.

In an effort to test the effect of the nudge, we randomly selected some reviewers to receive a treatment message in their initial invitation to review. The sample size of reviewers in each treatment condition and other descriptive statistics can be found in appendix A (apps. A–C are available online). Factors such as rank and gender were uncorrelated with placement in the treatment and control conditions, meaning that the randomization “worked.” In an effort to increase the likelihood that subjects received the treatment, the language in that message was inserted near the top of the invitation e-mail, before technical instructions for the review process, and reads: “SPPQ is embarking on a new initiative to ensure that all SPPQ papers include more thorough literature reviews. As part of this initiative, please include in your comments any relevant literature that the author(s) should be citing, including

citations to your own work. Please include this information in the confidential comments to the editor, and we will be sure to pass this along to the authors.”

We then counted number of citations per review, either in the comments to the author or the editor, although virtually all citations appeared in the comments to the author. We examined only the behavior of the reviewers in this experiment; we did not go back and determine whether the authors actually complied with the reviewers’ suggestions, although we think it is very likely they did because there is no cost at all to doing so and failure to comply could result in the author alienating a reviewer. By comparing the citation patterns of the reviews from the treatment group, who received the “nudge,” and the control group, who did not, we hoped to assess the effectiveness of such a nudge in increasing the number of citations to women and the number of self-citations from female reviewers.

THE REVIEW PROCESS

The process of selecting reviewers for each manuscript began with the editorial assistant receiving the manuscript and assigning it to one of the two journal editors and providing a list of potential reviewers. The editor then chose reviewers and informed the editorial assistant that the invitations are ready to be sent out. The editorial assistant then used an online random number generator to assign each reviewer to receive either a control invitation e-mail with no information or a treatment invitation that included the “nudge.”¹ The design was double blind: at no time did the editors know what treatment each reviewer received.

As reviews for each manuscript came in, the editorial assistant read the reviews and recorded the following data that we use in the analysis. Our dependent variables were whether the reviewers included additional citations, whether they included self-citations, and whether they included a citation that had a female author. Information about the variables and their measurement can be found in appendix B.

THE DATA

The universe of reviews we received from fall 2015 to the end of 2017 was a total of 488, 183 from women and 305 from men. Table 1 reports our power analysis, conducted using a set of simulations for a range of possible treatment proportions.²

1. In the treatment condition, we also randomized across which editor signed the invitation since one of the editors identifies male (Bonneau) and the other identifies female (Kanthak). There was no effect, and those results are not presented here.

2. Thanks to Christopher McConnell for his assistance in the construction of table 1.

Table 1. Power Analysis

Variable	α	$1 - \beta$	N_c	N_t	\hat{p}_c	$\beta > .2$	
						Min	Max
P(cites woman), woman	.05	.8	60	123	.23	.08	.42
P(cites woman), man	.05	.8	103	202	.19	.08	.33
P(self-cite), woman	.05	.8	60	123	.07	.04	.21
P(self-cite), man	.05	.8	103	202	.06	.03	.15

The last two columns in the table represent the range of values for \hat{p}_t (the mean probability of the reviewer in the treatment group citing) that would be detectable at 80% power, given the mean probability of the reviewer in the control group citing (\hat{p}_c) and the number of observations in the treatment (N_t) and control (N_c) groups.

The upshot of the power analysis is that our experiment is relatively underpowered, particularly given that relatively small differences in \hat{p}_t can translate into large differences in the number of works by women cited over the course of time. For example, in a journal that sees an increase in citations to women of 0.08 per review would see an increase over 400 reviews of 32 more citations to women. This experiment, however, represents what we believe is a first attempt to measure the effect the review process on gendered citation patterns, and so these results can provide us a preliminary answer to this question despite being somewhat underpowered.

A potential for a violation of the stable unit treatment values assumption (SUTVA) occurs if the same reviewer receives multiple invitations to review. It is important to note, however, that the control treatment is the language of the previous invitation, so in this sense, most of the treated had been subject to the control condition, if they had reviewed for the journal previously. A more worrisome SUTVA violation occurs if the subject had been treated and then randomly assigned to the control group in a subsequent invitation to review. This occurred 26 times in our 488 distinct reviews, all with a lag time of at least six months between treatments. Omitting those reviews does not change our results, but we retain them in the analysis, since our intent was to treat.

RESULTS

We were interested in the effects of the insertion of the message on both citation and self-citation patterns. Because some reviewers included citations in reviews even when not prompted, we can compare citation patterns of unprompted reviewers with those who received the prompt. Figure 1 assesses the probability that a reviewer cites an article with at least one

woman as an author (female reviewers who self-cite are included in this analysis). We divide the data by gender to assess any differences between male and female reviewers.

Figure 1 shows that female reviewers, in both the treatment and the control groups, are slightly more likely to cite at least one woman but that the treatment itself appears to have no significant effect on whether reviewers cite women. The bars in the figure are 90% confidence intervals and almost completely overlap. Note that about two-thirds of our reviewers received the treatment, which likely explains the smaller confidence intervals in the treatment condition. In all cases, about 20% of reviewers cite women, regardless of the gender of the reviewer or receipt of the “nudge” paragraph. Not shown here, the results also reveal no difference between the treatment and control groups for the probability of citing a man.

So the “nudge” does not appear to affect the probability of citing a woman, but perhaps it affects the number of women cited. In other words, the nudge may prompt those reviewers who were going to cite women (or men) in any event to cite more of them. In this case, the nudge would have an effect

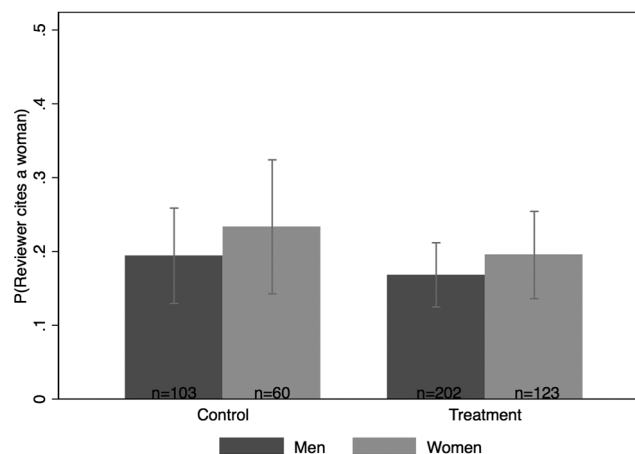


Figure 1. Probability of citing a female author by reviewer gender and experimental condition.

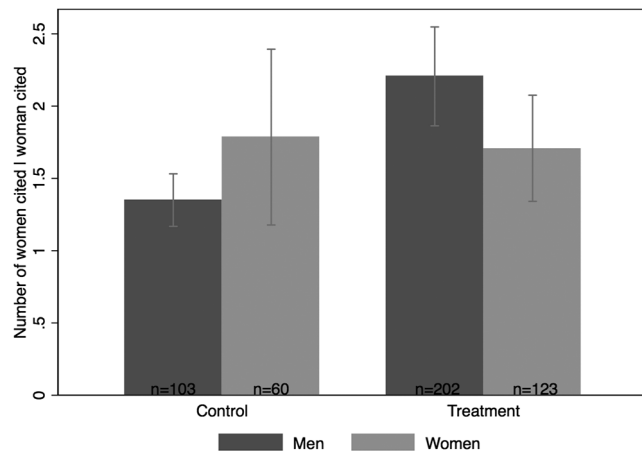


Figure 2. Number of women cited by reviewer gender and experimental condition.

on the number of women cited. We consider that possibility in figure 2.

Figure 2 shows that male, but not female, reviewers cite more women in the treatment condition than in the control condition, provided that the reviewer was already citing a woman. For men, the treatment prompted about one more citation to a woman author per review, a statistically significant effect. Notably, then, the net increase in citations to women in the treatment condition comes from men citing more women.

The treatment also invited subjects to include citations to their own relevant work. We present those results in figure 3.

The goal of the nudge was to provide women with an invitation to self-cite that might encourage them to do so. Figure 3 indicates that while the nudge seems to have an effect, it was not the intended one. Indeed, it is men, not women, who are more likely to self-cite when they are in the treatment condition. There is no significant gender difference in the control group, and in fact women are slightly more likely to self-cite than are men. Rather unexpectedly, then, the treatment condition creates a statistically significant self-citation gender gap where none exists in the control.³ Furthermore, the difference is based on a relatively modest decrease in women's self-citation patterns in the treatment condition, and a larger (6%) increase in men's self-citation patterns, although still insignificant at conventional levels. An actual increase in the self-citation gap, then, may be an artifact of these particular data, but the evidence appears strong that unlike in other contexts like electoral politics, inviting women to self-promote does not seem to mitigate gender differences in behavior. At the same time, this result could be because men are more likely than women to read the instructions and thus

3. The gender difference presented here is robust to inclusion in a logistic regression (available in app. C) controlling for academic rank.

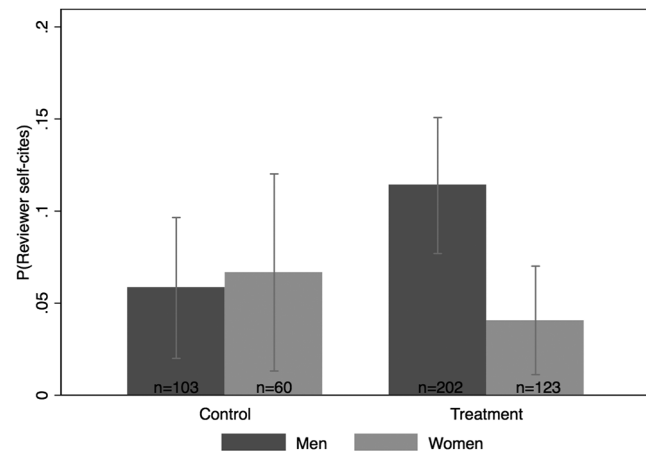


Figure 3. Probability of reviewer self-citing by reviewer gender and experimental condition.

more likely to receive the treatment. Although we know of no literature pointing to this conventional wisdom-defying prospect, we cannot rule it out.

DISCUSSION

Gender operates in the institution of the editorial review process because gender systematically affects patterns of citation. Furthermore, we show that changes to the institution can result in changes in the behavior of those who interact with that institution. The “editor’s nudge” works in the sense that it changes reviewers’ behavior, and it does so without an explicit mention of gender. Men in the treatment condition cite a greater number of women (see fig. 2), yet because men are more likely to self-cite (see fig. 3), inserting the nudge was not an effective means of mitigating the effect of gender in the review process. Blanket invitations to cite or self-cite more, then, are unlikely to mitigate the gender citation gap. However, the expected number of non-self-citations to women in the treatment group is 0.05 higher than the expected number in the control group. If this small effect is real, that translates to an expected 24 more citations to women in the 488 reviews included in the experiment. Certainly, more research is needed to parse these small, but potentially consequential, effects.

ACKNOWLEDGMENTS

The authors wish to thank the reviewers of *State Politics and Policy Quarterly*, the members of the American Political Science Association’s State Politics and Policy Section, and the reviewers and editors at *JOP* for helpful comments. Previous versions of this article were presented at the 2018 Institutions in Context: Gender Equality and Policy summer workshop at the University of Tampere and the 2018 meetings of the American Political Science Association.

REFERENCES

- Acker, Joan. 1992. "From Sex Roles to Gendered Institutions." *Contemporary Sociology* 21:565–69.
- Alper, Joe. 1993. "The Pipeline Is Leaking Women All the Way Along." *Science* 260 (5106): 409–41.
- Colgan, Jeff. 2017. "Gender Bias in International Relations Graduate Education? New Evidence from Syllabi." *PS: Political Science and Politics* 50:456–60.
- Fowler, James H., and Dag W. Aksnes. 2007. "Does Self-Citation Pay?" *Scientometrics* 72:427–37.
- Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus & Giroux.
- Kanthak, Kristin, and George A. Krause. 2012. *The Diversity Paradox: Political Parties, Legislatures, and the Organizational Foundations of Representation in America*. New York: Oxford University Press.
- Kerevel, Yann P., and Lonna Rae Atkeson. 2013. "Explaining the Marginalization of Women in Legislative Institutions." *Journal of Politics* 75:980–92.
- King, Molly M., Carl T. Bergstrom, Shelley J. Correll, Jennifer Jacquet, and Jevin D. West. 2017. "Men Set Their Own Cites High: Gender and Self-Citation across Fields and over Time." *Socius* 3:1–22.
- Larivière, Vincent, Ni Chaoqun, Yves Gingras, Blaise Cronin, and Cassidy R. Sugimoto. 2013. "Global Gender Disparities in Science." *Nature* 504 (7479): 211–13.
- Lawless, Jennifer L., and Richard L. Fox. 2005. *It Takes a Candidate: Why Women Don't Run for Office*. New York: Cambridge University Press.
- Maliniak, Daniel, Ryan Powers, and Barbara F. Walter. 2013. "The Gender Citation Gap in International Relations." *International Organization* 67:889–922.
- Mitchell, Sara McLaughlin, Samantha Lange, and Holly Brus. 2013. "Gendered Citation Patterns in IR Journals." *International Studies Perspectives* 14:485–92.
- Ostby, Gudrun, Havard Strand, Ragnhild Nordas, and Nils Petter Gleditsch. 2013. "Gender Gap or Gender Bias in Peace Research? Publication Patterns and Citation Rates for *Journal of Peace Research*, 1983–2008." *International Studies Perspectives* 14:493–506.
- Symonds, Matthew R. E., Neil J. Gemmill, Tamsin L. Braisher, Kylie L. Gorringer, and Mark A. Elgar. 2006. "Gender Differences in Publication Output: Towards an Unbiased Metric of Research Performance." *PLoS ONE* 1 (1): e127.
- Teele, Dawn Langan, and Kathleen Thelen. 2017. "Gender in the Journals: Publication Patterns in Political Science." *PS: Political Science and Politics* 50:433–47.