Gender (Im)balance in Citation Practices in Cognitive Neuroscience

Jacqueline M. Fulvio[1], Ileri Akinnola[2], and Bradley R. Postle[1]

Abstract

■ In the field of neuroscience, despite the fact that the proportion of peer-reviewed publications authored by women has increased in recent decades, the proportion of citations of women-led publications has not seen a commensurate increase: In five broad-scope journals, citations of papers first- and/or last-authored by women have been shown to be fewer than would be expected if gender was not a factor in citation decisions [Dworkin, J. D., Linn, K. A., Teich, E. G., Zurn, P., Shinohara, R. T., & Bassett, D. S. The extent and drivers of gender imbalance in neuroscience reference lists. *Nature Neuroscience*, *23*, 918–926, 2020]. Given the important implications that such underrepresentation may have on the careers of women researchers, it is important to determine whether this same trend is true in subdisciplines of the field, where interventions might be more targeted. Here, we report the results of an extension of the analyses carried out by Dworkin et al. (2020) to citation patterns in the *Journal of Cognitive Neuroscience*. The results indicate that the underrepresentation of women-led publications in reference sections is also characteristic of papers published in *Journal of Cognitive Neuroscience* over the past decade. Furthermore, this pattern of citation imbalances is present regardless of author gender, implicating systemic factors. These results contribute to the growing body of evidence that intentional action is needed to address inequities in the way that we carry out and communicate our science. ■

INTRODUCTION

The public dissemination of research findings is critical for the advancement of any field of scientific inquiry. Similarly, evidence of impactful publication in peer-reviewed journals is critical for a researcher's advancement in their field. For example, citation-based metrics such as impact factors, the h-index (Hirsch, 2005), and the i10-index (Connor, 2011) contribute to the evaluation of one's scholarly "worth" (Fairhall & Marder, 2020). It is problematic, therefore, that citation-based metrics of scholarship in neuroscience show a gender bias. A recent study evaluating citation practices in five broad-scope neuroscience journals—*Brain*, *Journal of Neuroscience*, *Nature Neuroscience*, *Neuroimage*, and *Neuron*—demonstrated over-citation of papers published by men as first and last authors compared with the rate expected if gender did not play a role in citation choices, whereas papers published by a woman in the first- and/or last-author position have been undercited (Dworkin et al., 2020).

The findings of Dworkin et al. (2020) are a cautionary tale for fields grappling with gender disparities because bias in citation practices may limit the advancement of individual researchers as well as the advancement of their approaches and ideas. Quantification and dissemination of evidence of such biases is an important first step toward developing more equitable practices. Here, we sought to determine whether the gender imbalance in citation practices reported for broad-scope neuroscience journals is also characteristic of the *Journal of Cognitive Neuroscience* (*JoCN*), the flagship journal of this subdiscipline of neuroscience.

METHODS

We applied the methodological approach used by Dworkin et al. (2020), using their open-source R code (osf.io/h79g8/). Where necessary, we modified the code to support the *JoCN*-specific analysis.

Data Acquisition

The data for the analysis were obtained from the Web of Science Web site (www.webofknowledge.com/). Metadata for the 2106 research papers and review papers published in *JoCN* from January 2009 to July 2020 were downloaded. We note that metadata for *JoCN* papers are available dating back to 1995, but metadata from before 2009 contain author initials rather than full first names, with the latter being necessary for the analysis. In addition, preprocessed metadata from broad-scope neuroscience journals—*Brain*, *Journal of Neuroscience*, *Nature Neuroscience*, *Neuroimage*, and *Neuron* (from here forward, the "broad-scope journals")—were obtained from Jordan Dworkin with permission for use in the analysis described here.

[1]University of Wisconsin–Madison, [2]University of Maryland, Baltimore County

Gender Category Assignment

Gender Category Assignment of Papers Published in JoCN

We first extracted the names of each author of each publication from the metadata into an array with the general format "last name, first name; last name, first name." We then implemented the algorithm used by Dworkin et al. (2020) to disambiguate authors with different versions of their names across papers (such as instances with and without middle initials or with and without nicknames). In brief, the algorithm matches entries first by last name and then by the same first and/or middle name or initials and assigns the most common first name variant to all instances.

Next, the first names of the first and last authors of each paper were assigned a probability of belonging to someone self-identifying with either of two gender labels—"man" or "woman." (Note that the a priori assumption that gender identification is a binary variable is not valid but was necessitated by limitations of our method.) First, each name was queried within the Social Security Administration baby name data set, which assigns labels based on the sex assigned at birth. If the name was found, the probabilities of that name belonging to someone self-identifying as "man" and "woman" were returned. If the name was not found, it was submitted to Gender API (gender-api.com/) for probability assignment. Gender API includes approximately 815,000 unique first names from 189 countries and assigns labels based on a combination of the sex assigned at birth and genders detected in social media profiles. Using the same criteria as Dion, Sumner, and Mitchell (2018) and Dworkin et al. (2020), we assigned a gender label to each author if their name had a probability ≥ .70 of belonging to someone of either gender. The author's gender label was manually assigned in instances where Gender API returned a probability < .70 if the author had publicly available pronouns (e.g., on their personal or university faculty Web sites) following Dworkin et al. (2020).

The 2106 research papers published in *JoCN* were then assigned an authorship gender category ("man/man" [MM], "woman/man" [WM], "man/woman" [MW], or "woman/woman" [WW]) based on the assigned gender labels of the first and last authors. Upon completion of this step, ~9% of the papers had incomplete authorship gender category designations because of single authorship, first name initials, poor formatting because of incorrect parsing of middle initials, or other formatting problems that arose during metadata extraction that impeded the name query steps. In these cases, we performed a manual correction step to hand-code the category designations (Dion et al., 2018), sometimes entailing visits to the paper's page on the journal's Web site or the author's Web site. Single-authored papers were assigned "MM" or "WW" according to the assigned gender label of the author.

We note that the automated, probabilistic nature of gender assignment method may be subject to sources of bias in the results. For example, bias could arise if missing metadata were skewed across gender lines, or the Gender API queries tended to return a higher proportion of equivocal results on names for which the ground truth label is "woman," for example. However, as indicated above, these and other impediments to gender assignment occurred for a small proportion (~9% of papers), to which we then applied manual correction. Furthermore, Dworkin et al. (2020) carried out an independent test of the method and found high accuracy (>90%) on both individual author and paper gender category assignments.

Gender Category Assignment of Papers Cited in Papers Published in JoCN

A final critical preprocessing step was to assign authorship gender categories to the papers cited in the 2106 *JoCN* papers. We first extracted each *JoCN* paper's citation list from the metadata. Importantly, although the extracted citation lists did not contain author first names (only initials), the metadata did include the Digital Object Identifier (DOI) associated with each citation. These DOIs allowed us to use the preprocessed data from six journals—the five broad-scope journals plus *JoCN*—as a lookup table, in which we attempted to match each citation DOI with a DOI of a paper published in one of the six journals. If a match was found, we assigned that cited paper to the authorship gender category assigned to the original paper (from the previous step for *JoCN* and from analogous data for the five broad-scope journals analyzed by Dworkin et al. [2020]). If citation list metadata for a particular paper were incomplete, the matching algorithm attempted to match any available DOIs in the list; if the metadata were missing, the matching algorithm proceeded with the next paper. No manual correction of citation list data was performed. In all, citations from 2069 of the 2106 papers (98.3%) published in *JoCN* from 2009 to July 2020 were matched with other papers from *JoCN* and the five broad-scope journals, with 20.5% of the total references being assigned an authorship gender category including self-citations. Self-citations constituted 28.1% of categorized citations. As described below, self-citations were removed when computing Gender Balance Citation Indices, which therefore were based on 14.7% of the total references cited by *JoCN* papers published from 2009 to July 2020.

Categorical Gender Quantification

Quantification of JoCN Authorship

The gender balance of authorship in *JoCN* was quantified in three ways: collapsing across the January 2009 to July 2020 time frame, broken out into each publication month, and as the cumulative portion of the overall time frame leading up to each publication month and year. This latter set of proportions served as the base rate (i.e., the "expected" proportions) for the Gender Citation Balance Index calculations described below.

Computation of Gender Citation Balance Indices

Our primary goal was to determine how the gender proportions in the reference lists of papers published in *JoCN* corresponded to the gender proportions of *JoCN* authorship. Following Dworkin et al. (2020), we removed self-citations (i.e., cited papers for which either the first or last author was the either first or last author on the citing paper) to remove effects of gender differences in self-citation behaviors (King, Bergstrom, Correll, Jacquet, & West, 2017) and focused instead on authors' citation of other researchers in the field. For each of the 2069 papers that cited other papers from *JoCN* and the five broad-scope journals, we computed the proportion of those citations assigned to each of the four gender citation categories, which were designated the "observed" proportions.

We computed Gender Citation Balance Indices for each of the four gender citation categories as

$$\textit{Gender Citation Balance Index} = \frac{\textit{observed prorportion} - \textit{expected proportion}}{\textit{expected proportion}} \quad (1)$$

Thus, positive values corresponded to more frequent citations of the category than expected, and negative values corresponded to less frequent citations of the category than expected. (Note that, because we did not have *JoCN* publication data before January 2009, we used the *JoCN* authorship in each category during January 2009 for the expected rates of papers published in January 2009.) Finally, we bootstrapped the 95% confidence interval for each category using 1000 iterations of random sampling with replacement from the 2069 papers. For each iteration, we determined the Gender Citation Balance Index for each category, and the 2.5th and 97.5th percentiles corresponded to the lower and upper bounds of the confidence interval for that category, respectively.

RESULTS

The categorical gender breakdown in *JoCN* authorship has been relatively stable from 2009 to mid-2020, with an increase in WW-authored publications in the most recent years (Figure 1). Overall, 40.8% of *JoCN* papers published during this timeframe were MM, with the remaining 59.2% having at least one woman in the first or last author positions (i.e., W ∪ W). Interestingly, this proportion of W ∪ W papers in *JoCN* is considerably larger than the average of 44.7% W ∪ W-authored papers for the five broad-scope neuroscience journals that are also included in our analyses.

The Gender Citation Balance Indices of papers published in *JoCN* reveal an overcitation of MM papers and an undercitation of WM, MW, and WW papers (Figure 2). During the 2009 to July 2020 timeframe, MM papers had a base rate of 40.8%, but they accounted for 57.9% of categorized citations when self-citations were removed. WM papers had a base rate of 33.5% but accounted for 24.3% of categorized citations. MW papers had a base rate of 10.8% but accounted for 10.2% of categorized citations. Finally, WW papers had a base rate of 14.9% but accounted for 7.6% of categorized citations.

Importantly, this qualitative pattern is observed across author "gender subgroupings" when papers are broken out by author gender category, although MW papers have a positive Gender Citation Balance Index for papers from the MW and WW gender subgroups (Figure 3).

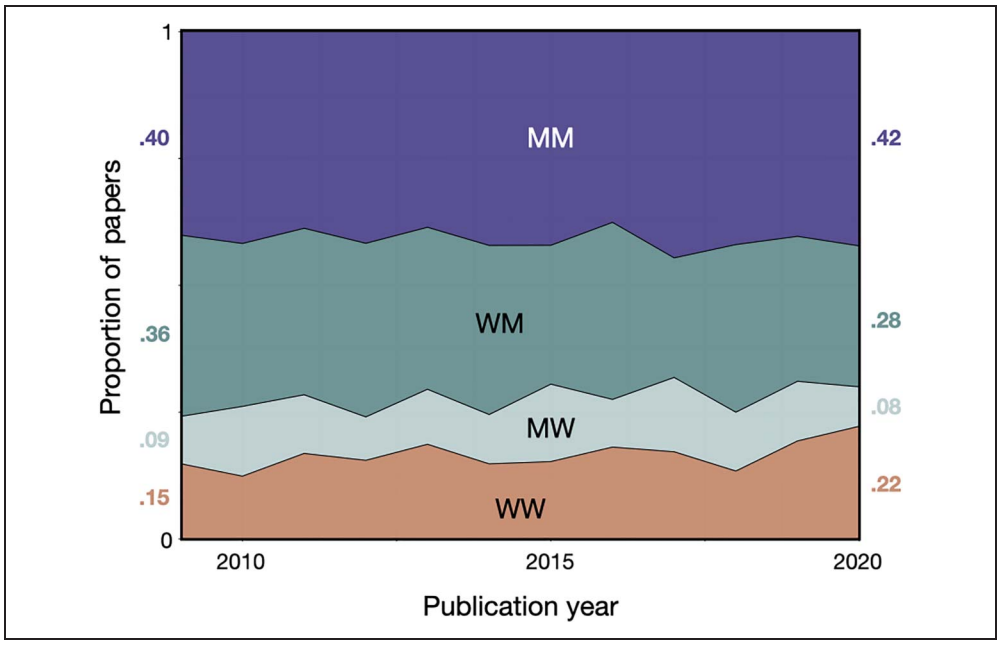

Figure 1. Gender breakdown in *JoCN* authorship from 2009 to 2020. Proportion of *JoCN* papers assigned to four categories: men as the first and last author (MM; purple), women as the first author and men as the last author (WM; darker green), men as the first author and women as the last author (MW; lighter green), and women as both the first and last author (WW; salmon). For ease of comparison across time, the proportions of each category are indicated for 2009 (left) and 2020 (right).

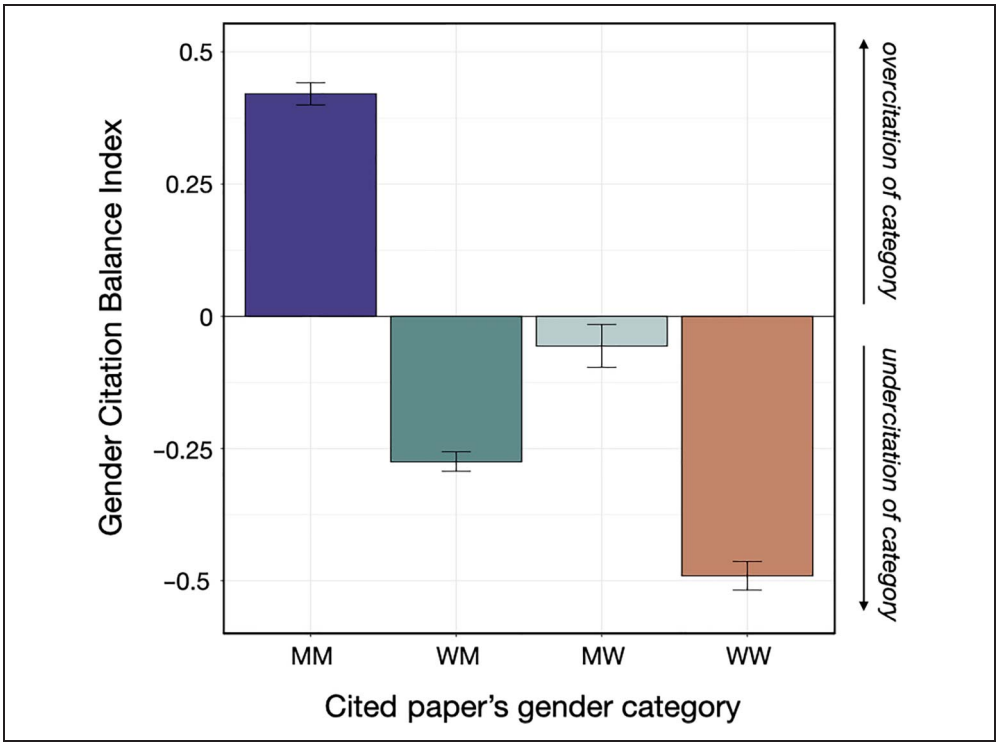

Figure 2. Gender Citation Balance Indices for the four gender categories of peer-reviewed papers published in *JoCN*. Error bars correspond to bootstrapped 95% confidence intervals.

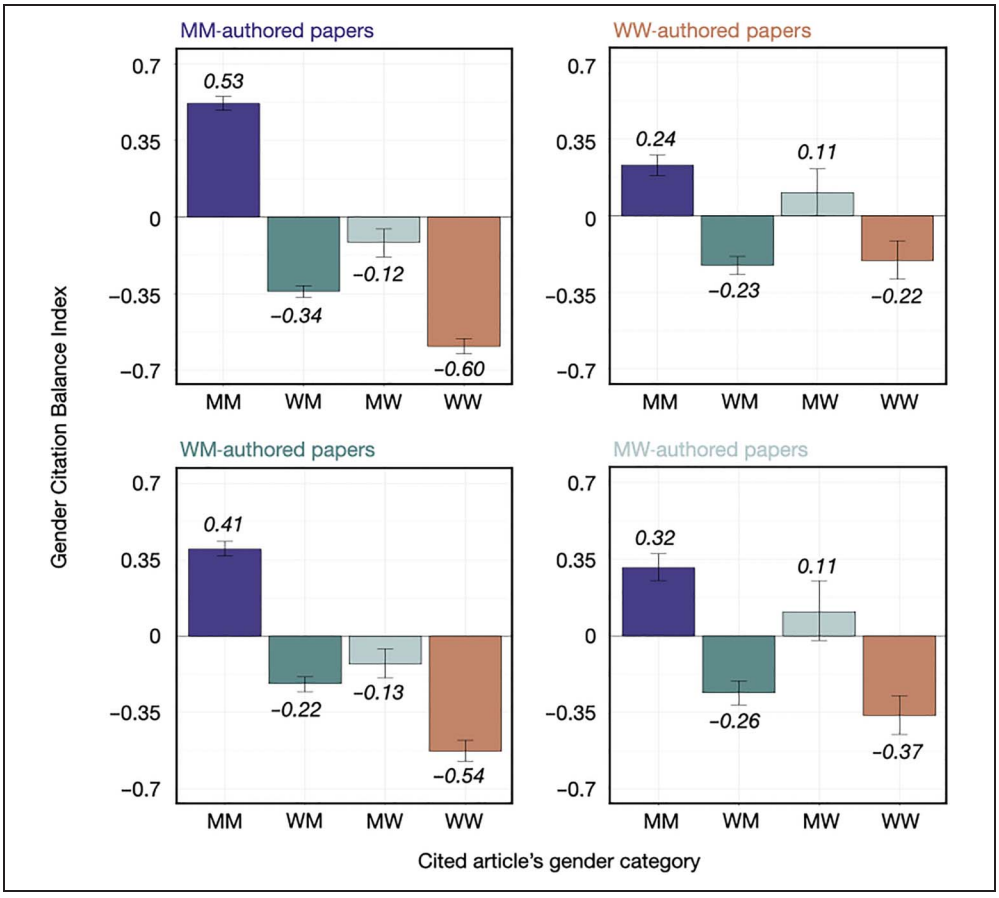

Figure 3. Gender Citation Balance Indices for peer-reviewed papers published in *JoCN*, broken down by citing papers' gender category. Error bars correspond to bootstrapped 95% confidence intervals.

DISCUSSION

We measured the degree to which the categorical gender balance of papers cited in *JoCN* from 2009 to 2020 reflected the gender balance of the authorship of the journal during that timeframe. The results indicate that papers authored by men as the first and last authors have been overcited compared with what would be expected based on the number of papers published by the journal that were authored by "MM" teams. By contrast, papers authored by teams with at least one woman in the first- and/or last-author position have been undercited.

These findings indicate that the gender imbalance in citation practices that was previously reported for broad-scope neuroscience journals (Dworkin et al., 2020) extends to the subfield of cognitive neuroscience. The fact that this pattern of imbalance is present in *JoCN* papers published by each of the four gender-defined groups that we have considered here (MM, WM, MW, and WW) indicates that this imbalance results, at least in part, from systemic factors at play in the field overall.

Limitations

There are caveats to bear in mind when interpreting these data. A fundamental one is that it assumes that all scientists self-identify within a gender-binary framework. This is, of course, not true and results in two limitations. First, it introduces error into the estimates for authors who do identify as female or male. Second, it highlights that this study does not speak to inequities faced by noncisgender individuals and by members of the LGTBQ+ community. Methodologically, our reliance on data from the five broad-scope neuroscience journals may have somewhat skewed our estimates of Gender Citation Balance Index toward positive values for MM papers because the base rate of authorship in those journals is more heavily weighted toward MM than is the base rate for *JoCN*. Working in the opposite direction, however, is the fact that our method included removing self-citations. Because men self-cite at a higher rate than do women (King et al., 2017), including self-citations would be expected to push the Gender Citation Balance Index for MM papers further in the positive direction. (It is worthy of note, however, that it is unlikely that readers of journal papers somehow selectively "remove" the influence of self-citations from their internal model of gender balances in scientific publishing.)

Conclusion

In carrying out these analyses, we deliberately limited ourselves to the "simple first step" (cf. Dworkin et al., 2020) of quantifying and describing this phenomenon. Although we lack the expertise to propose specific interventions that may encourage prosocial behavior, it is our hope that this work contributes, in some modest way, to social norm messaging (Murrar, Campbell, & Brauer, 2020) about the need to address inequities in the way that we carry out and communicate our science.

Acknowledgments

This work was supported by NIH grant MH064498. I. A. received support from National Science Foundation grant 1757785.

Reprint requests should be sent to Jacqueline M. Fulvio, Department of Psychology, University of Wisconsin–Madison, 1202 W. Johnson St., Madison, WI 53706, or via e-mail: jacqueline.fulvio@wisc.edu.

Author Contributions

Jacqueline Fulvio: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Software; Supervision; Validation; Visualization; Writing - Original Draft; Writing - Review & Editing. Ileri Akinnola: Data curation; Formal analysis; Investigation; Methodology; Visualization. Brad Postle: Conceptualization; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Validation; Writing - Review & Editing.

Funding Information

Jacqueline Fulvio, National Science Foundation (http://dx.doi.org/10.13039/100000001), Grant number: 1757785. Jacqueline Fulvio, National Institutes of Health (http://dx.doi.org/10.13039/100000002), Grant number: MH064498.

REFERENCES

Connor, J. (2011). Google Scholar citations open to all. *Google Scholar Blog*. https://scholar.googleblog.com/2011/11/google-scholar-citations-open-to-all.html.

Dion, M. L., Sumner, J. L., & Mitchell, S. M. (2018). Gendered citation patterns across political science and social science methodology fields. *Political Analysis*, *26*, 312–327. **DOI:** https://doi.org/10.1017/pan.2018.12

Dworkin, J. D., Linn, K. A., Teich, E. G., Zurn, P., Shinohara, R. T., & Bassett, D. S. (2020). The extent and drivers of gender imbalance in neuroscience reference lists. *Nature Neuroscience*, *23*, 918–926. **DOI:** https://doi.org/10.1038/s41593-020-0658-y, **PMID:** 32561883

Fairhall, A. L., & Marder, E. (2020). Acknowledging female voices. *Nature Neuroscience*, *23*, 904–905. **DOI:** https://doi.org/10.1038/s41593-020-0667-x, **PMID:** 32591766

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences, U.S.A.*, *102*, 16569–16572. **DOI:** https://doi.org/10.1073/pnas.0507655102, **PMID:** 16275915, **PMCID:** PMC1283832

King, M. M., Bergstrom, C. T., Correll, S. J., Jacquet, J., & West, J. D. (2017). Men set their own cites high: Gender and self-citation across fields and over time. *Socius: Sociological Research for a Dynamic World*, *3*, 237802311773890. **DOI:** https://doi.org/10.1177/2378023117738903

Murrar, S., Campbell, M. R., & Brauer, M. (2020). Exposure to peers' pro-diversity attitudes increases inclusion and reduces the achievement gap. *Nature Human Behaviour*, *4*, 889–897. **DOI:** https://doi.org/10.1038/s41562-020-0899-5, **PMID:** 32601460