

Citation Diversity Statement

Recent work in several fields of science has identified a bias in citation practices such that papers from women and other minority scholars are under-cited relative to other papers (Caplar et al., 2017; Chakravartty et al., 2018; Dion et al., 2018; Dworkin et al., 2020; Maliniak et al., 2013; Mitchell et al., 2013; Trepte & Loths, 2020; Wang et al., 2020). In this work, we sought to proactively select references that reflect the diversity of the field in thought, form of contribution, gender, race, ethnicity, and other factors.

In order to ascertain the diversity of the citations in this manuscript, we relied on the procedure outlined by Zhou and colleagues (2021).

First, we obtained predictions for the gender of the first and last author of each reference cited in this manuscript using the Gender API (Dworkin et al., 2020; Zhou et al., 2021). Note that this method limited in that a) names, pronouns, and social media profiles used to construct the databases may not, in every case, be indicative of gender identity and b) it cannot account for intersex, non-binary, or transgender people. By this measure (and excluding self-citations to the first and last authors of our current paper), 14.29% of references in this paper have a woman as first and last author, 6.49% have a man first author and a woman last author, 14.78% have a woman first author and a man last author, and 64.44% have a man as first and last author.

Second, we obtained the predicted racial/ethnic category of the first and last author of each reference using extant databases that store the probability of a first and last name being carried by an author of color (Ambekar et al., 2009; Sood & Laohaprapanon, 2018). This method is limited in that a) names and Florida Voter Data used to make the predictions may not be indicative of racial/ethnic identity, and b) it cannot account for Indigenous and mixed-race authors, or those who may face differential biases due to the ambiguous racialization or ethnicization of their names. By this measure (and excluding self-citations), 11.39% of our references have an author of color as both the first and last author, 18.15% have a first author who is white and a last author of color, 10.29% have a first author of color and a last author who is white, and 60.17% have a first and last author who is white.

The limitations noted above imply that the outcome measures reported herein may contain inaccuracies, or even systematic biases of their own. We look forward to future work that could help us to develop more accurate and sensitive measures of biases in citation practices, and to better understand how to support equitable practices in science.

References

- Ambekar, A., Ward, C., Mohammed, J., Male, S., & Skiena, S. (2009). Name-ethnicity classification from open sources. *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 49–58.
<https://doi.org/10.1145/1557019.1557032>
- Caplar, N., Tacchella, S., & Birrer, S. (2017). Quantitative evaluation of gender bias in astronomical publications from citation counts. *Nature Astronomy*, 1(6), 1–5.
<https://doi.org/10.1038/s41550-017-0141>
- Chakravarty, P., Kuo, R., Grubbs, V., & McIlwain, C. (2018). #CommunicationSoWhite. *Journal of Communication*, 68(2), 254–266. <https://doi.org/10.1093/joc/jqy003>
- Dion, M. L., Sumner, J. L., & Mitchell, S. M. (2018). Gendered citation patterns across political science and social science methodology fields. *Political Analysis*, 26(3), 312–327.
<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(8), 918–926. <https://doi.org/10.1038/s41593-020-0658-y>
- Maliniak, D., Powers, R., & Walter, B. F. (2013). The gender citation gap in international relations. *International Organization*, 67(4), 889–922.
<https://doi.org/10.1017/S0020818313000209>
- Mitchell, S. M., Lange, S., & Brus, H. (2013). Gendered citation patterns in international relations journals. *International Studies Perspectives*, 14(4), 485–492.
<https://doi.org/10.1111/insp.12026>
- Sood, G., & Laohaprapanon, S. (2018). *Predicting Race and Ethnicity From the Sequence of Characters in a Name*. <https://arxiv.org/abs/1805.02109v1>
- Trepte, S., & Loths, L. (2020). National and gender diversity in communication: A content analysis of six journals between 2006 and 2016. *Annals of the International Communication Association*, 44(4), 289–311.
<https://doi.org/10.1080/23808985.2020.1804434>
- Wang, X., Dworkin, J. D., Zhou, D., Stiso, J., Falk, E. B., Bassett, D. S., Zurn, P., & Lydon-Staley, D. M. (2021). Gendered citation practices in the field of communication. *Annals of the International Communication Association*. Advance online publication.
<https://doi.org/10.1080/23808985.2021.1960180>
- Zhou, D., Stiso, J., Cornblath, E., Teich, E., Sizemore-Blevins, A., Oudyk, K., Michael, C., Virtualmario, C., & Camp, C. (2021). Gender Diversity Statement and Code Notebook (1.1.1) [Computer software]. <https://doi.org/10.5281/zenodo.3672109>

Figures

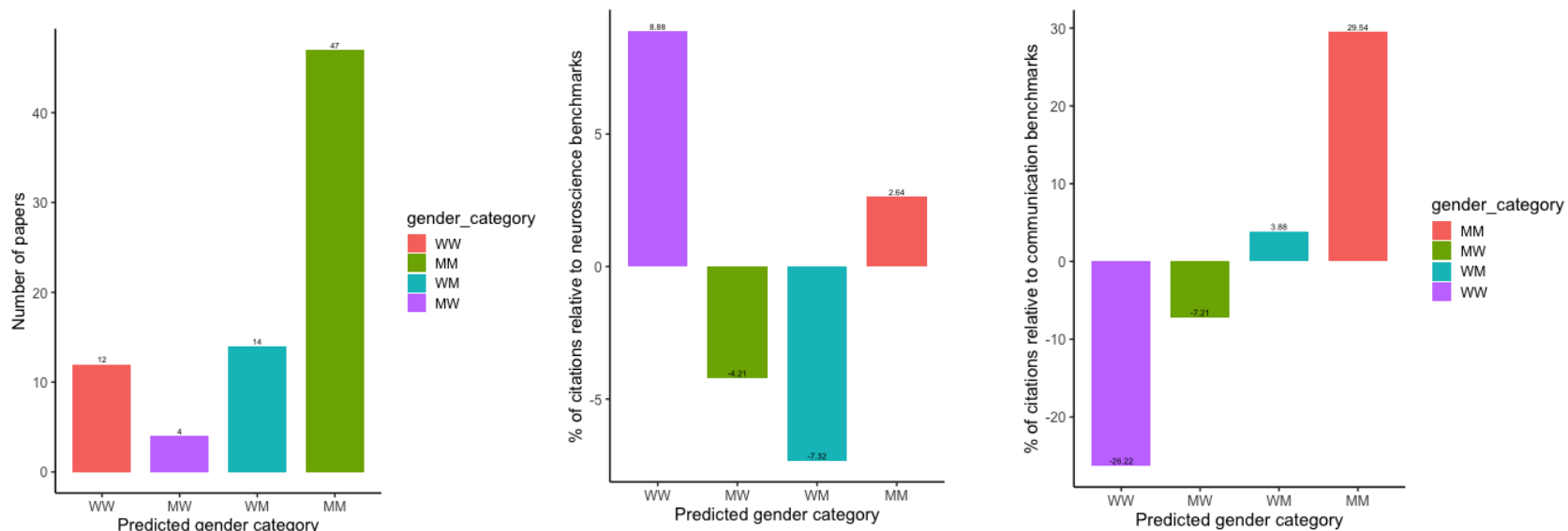


Figure S18: Histogram of Citations by Author Category

Distribution of citations by author category (first author/last author) for woman/woman (WW), man/woman (MW), woman/man (WM), man/man (MM), and unknown/unknown (UU) author teams. (Left) The raw frequency counts for citations by author category. (Right) The percent of over/under representation by author category. Neuroscience benchmarks are drawn from Dworkin and colleagues (2020), and communication benchmarks are drawn from Wang and colleagues (2020).