Publishing While Female: are Women Held to Higher Standards Regardless of Native Language?*

Sindhu Priya Mallavarapu Jena Shah

February 15, 2024

This research examines the effects of English-proficiency on the standards applied to academic publishing. We are building upon previous research suggesting that women may be held to higher standards in scholarly writing and investigate if there are differences for non-native English speakers. Utilizing data from CSV datasets, we analyzed the Flesch readability scores of papers written by native and non-native English speakers, filtered by gender. Our findings revealed no significant difference in the readability scores based on language proficiency. Albeit, consistent with prior research, we observed gender disparities with women facing higher standards than men. These results highlight the importance of addressing biases in academic publishing and advocating for equality in the area. Future research should explore the reasons behind gender disparities in publishing standards to help identify strategies which can promote inclusivity in scholarly publishing.

1 Introduction

The process of academic publishing plays an important role in presenting research findings and advancing our knowledge in the scholarly communities. However, there have been concerns raised with the potential biases in evaluating research papers, particularly based on factors such as language proficiency and gender of the author. Previous studies have suggested that females may face higher standards in academic publishing, as shown by the differences in the readability scores compared to their male counterparts. Additionally, the impact of language proficiency–specifically for non-native English speakers–on publishing standards remains an area of interest and concern for us.

*Code and data are available at: https://github.com/MSindhuPriya/publish_while_female. The replication is available at :

Erin Hengel conducted research which showed that female written papers tend to be better than ones written by males and they also take longer to review (Hengel (2022)). In her paper, she assesses different factors other than So to add onto Hengel's research, our study investigates the connection between language proficiency, gender, and readability–which we relate to evaluation standards in academic research. We aim to assess whether non-native English speakers are subject to higher standards in academic publishing and whether gender differences persist even in this context. In particular we want to measure and compare the mean readability scores of native English speaking female publishers, non-native English speaking female publishers, native English speaking male publishers, and non-native English speaking male publishers.Thus, the estimand in this paper is the mean readability scores of these our demographics of publishers. By analyzing the readability scores of papers authored by native and non-native English speakers and filtering it further through gender, we aim to provide insights into potential biases in the evaluation of scholarly work. Understanding how this works can be beneficial as it can help with encouraging equity, ensuring fairness and diversity with the evaluation of future papers being published.

In the following Data section this paper will introduce the data that was used and the work that was done to get the data ready for analysis.This section will also include any limitations in measurement of the data. There, an initial glance at the variables used will also be provided. In the Results section, we show that regardless of native language, female publishers tend to have better written papers than male publishers. In the discussion we talk about what these results can mean for the world and future steps that can be taken to further explore the relationship between readability and gender. This section will also cover limitations of the paper. The conclusion will wrap up the entire paper and summarize findings. All the analysis for this paper is done through R (R Core Team (2022)). Additionally, we also use the help of various libraries in R like `tidyverse` (Wickham et al. (2019)), `janitor` (Firke (2023)), `knitr` (Xie (2014)), `ggplot2` (Wickham (2016)), and `dplyr` (Wickham et al. (2023)).

2 Data

2.1 Overview

The replication was done on the paper titled "Publishing While Female: Are Women Held to Higher Standards? Evidence from Peer Review" (Hengel (2022)). This was taken from The Economic Journal published by the Oxford Academy. This original paper argues that female publishers are held to higher standards than male publishers. It presents four facts in the paper, namely, (i) female papers are 1%-6% better written than papers by male publishers, (ii) the gap between the readability of these papers widens during peer review, (iii) while men do not improve their writing as time goes on, women do, and (iv) papers written by female publishers are put under review for much longer than papers written by male publishers. The reproduction of this paper further extends these insights by exploring whether being a native english speaker or not changes the conclusions drawn in the original paper. Specifically, this

reproduction explores whether native and non-native english speaking female publishers have the same readability scores.

In this paper, the term "better written" refers to papers that are easier to read and understand. This can be measured using various metrics, but this paper uses Flesch scores to measure the readability of papers. Further, "readability" is used in this paper to be synonymous with how easy it is to read and understand a paper. Moreover, "readability scores" refer to the metrics that measure "readability". We chose to use Flesch scores because they directly measured readability in that higher Flesch scores indicate a more readable paper. The other scores such as Flesch-Kincaid or Gunning fog generate the grade level estimating the minimum level of education needed to understand the paper. Since Flesch scores more directly measure what the paper is trying to explore it was chosen. Moreover, other readability scores were inversely proportional to readability which contributed another reason as to why they were rejected.

The data for this paper came from the replication package associated with the paper that was used for replication. This package is available through Zenodo [add citation]. For the purposes of the replication and the reproduction, we only needed some of the datasets available in the replication package. In particular the `AuthorCorr.csv`, `Author.csv`, `primary_jel.csv`, and `ReadStat.csv` were used. Other datasets contained information pertaining to the number of children the researchers have and the year of their birth, all of which are outside of the scope of this paper, so they were not used. Merging the abovefiles into one gave us a singular csv with the variables `article_id` a unique value assigned to each individual article, `author_id` a unique number assigned to each individual author, `stat_name` the name of the readability metric used (in our case we filtered for the flesch_score metric when merging), `stat_value` the actual value of the readability metric, `author_name`, `sex` the sex of the author, `native_language` a variable indicating whether or not their native language is English, and `jel` a variable that tells you the category of the paper (ex. General, Methodology, etc.).

2.2 Measurement

In our study, we used and analyzed the readability scores. These scores did not account for some factors such as grammar, which can be a potential bias in the analysis. Due to this, the overall results may have some errors since the inclusion of grammar when grading readability may have significantly changed the results.

3 Results

We combined the CSV files provided in the original research. We then proceeded to find and compare the central tendencies of the Flesch readability scores of native English speaking females, native English speaking males, non-native English speaking females, and non-native English speaking males. The results obtained are shown in the table below.

	Native Female	Native Male	Non-native Female	Non-native Male
Maximum	79.26824	83.18090	70.63400	79.75871
Mean	42.21659	40.55975	41.03498	40.11318
Median	42.77569	41.21493	41.18377	40.78191
Standard Deviation	11.38406	11.97531	11.21791	12.21947

As seen from the table above (Table 1: Summary Statistics), native English speaking women have the highest average readability scores on their papers, with non-native English speaking women having the second highest average scores. The next highest readability scores are those of native English speaking men. According to the data, men who are non-native English speakers have the lowest average readability scores for their papers. Therefore, it can be concluded from the data that regardless of native language, female publishers overall write papers that are easier to read and understand than their male counterparts.

4 Discussion

Our investigation on the influence of English proficiency regarding the standards applied to publishing has provided insightful results. Contrary to our initial expectations, the analysis did not reveal significant differences in the readability scores between papers authored by native and non-native English speakers. This finding challenges the hypothesis of non-native speakers being held to higher standards in academic publishing.

However, our study verifies previous research indicating there are disparities in gender within the publishing standards. Despite the author's language proficiency, females have higher Flesch readability scores compared to men, ultimately showing that they may be held to a higher standard. The results show us the systemic biases that unequally affect female authors. Although we did not find evidence of language-based biases, the limitations must be acknowledged. There are factors beyond readability scores which may still influence the quality of scholarly work, such as language fluency, cultural nuances, and writing style. Future research in this topic should consider an approach which examines the relationships of language proficiency, gender, and other demographic factors.

5 Conclusion

In conclusion, our study contributes to the discussion of biases in academic publishing through examining the impact of language proficiency and gender on publishing standards. While we found no significant difference in readability scores between native and non-native English speakers, gender differences exist with females being held to higher standards. Moving forward, it is necessary to address these biases and strive for inclusivity when publishing research. The efforts made to promote equality should consist of not only language proficiency, but also

gender, race, ethnicity, and other measures of diversity. Through encouraging a more equality-based environment, we can ensure that all scholars have equal opportunities to contribute to the advancement of knowledge.

References

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Hengel, Erin. 2022. "Publishing While Female: Are Women Held to Higher Standards? Evidence from Peer Review." *The Economic Journal* 132 (648): 2951–91. https://doi.org/10.1093/ej/ueac032.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Xie, Yihui. 2014. *Knitr: A Comprehensive Tool for Reproducible Research in r*. Chapman; Hall/CRC.