Persistent Underrepresentation of Women's Science in High Profile Journals

Yiqin Alicia Shen, Jason M. Webster, Yuichi Shoda, and Ione Fine
Department of Psychology, University of Washington

Abstract

Past research has demonstrated an under-representation of female editors and reviewers in top scientific journals, but very few studies have examined the representation of women authors within original research articles. We collected research article publication records from 15 high-profile multidisciplinary and neuroscience journals for 2005-2017 and analyzed the representation of women over time, as well as its relationship with journal impact factor. We found that 1) Women authors have been persistently underrepresented in high-profile journals. This under-representation has persisted over more than a decade, with glacial improvement over time. 2) The percent of female first and last authors is negatively associated with a journal's impact factor. Since publishing in high-profile journals is a gateway to academic success, this underrepresentation of women may contribute to the lack of women at the top of the scientific academic ladder.

Introduction

It has long been known that female representation decreases at every stage of the scientific career (Advisors, 2017). Take neuroscience as an example, in the year 2016, over 55% of graduate students were female, however, only 45% of postdoc, 32% of faculty were female.

The reason behind such gender disparity is complex (Shen, 2013). One potential problem that has gathered increasing research interest is the gender discrepancy in scientific publication. For example, a series of prominent articles indicated that women are underrepresented as invited authors of *Nature* (Conley, 2005; "Gender imbalance in science journals is still pervasive," 2017; "Nature's sexism," 2012).

While commissioned opinion pieces are important and changes in author recruiting can be directly influenced by journal policies, one of the most influential functions of high profile journals such as *Nature* is to disseminate high-profile original findings. An initial small scale analysis to examine gender disparities in research articles across two 3-month periods in 2006 and 2016 in Nature Neuroscience ("Promoting diversity in neuroscience," 2018) found only a 1% increase in the number of female corresponding authors.

In the current research, we extend this work by using data mining techniques to examine the proportion of female first and last authors for all research articles published between 2005 and 2017 across a wide range of high-profile journals that publish neuroscience research. Here, we focus on two aspects: First, we show that, even within the highly selective group of journals, there was negative relationship between journal impact-factor and proportion of female first and last authors. Second, we show that the lack of representation of female authors has remained dispiritingly unchanged in most journals over the last 13 years.

Methods

The full details and code for data acquisition, processing, and analysis are given in the Github Repo (https://github.com/VisCog/Women-in-High-Profile-Journals). Here we describe an overview of our approach.

Data Acquisition

We downloaded metadata associated with all papers published from 2005 to 2017 from the PubMed's MEDLINE database ("MEDLINE/PubMed Data," 2017). We then subsetted to focus on research articles in those journals. We did this by excluding articles without an abstract.

In order to focus on high profile journals, we selected 15 journals to include based on the 2016 impact factors from the Thomson Reuters InCite Journal Citation Report (Analytics, 2016). Journals which focused on a particular aspect of neuroscience (e.g. EMBO, Stroke) were excluded. This resulted in a list that included both non-specialized multidisciplinary journals (Nature, Science, Proceedings of National Academy of Science), and top non-specialized journals in the field of neuroscience (Nature Review Neuroscience, Nature Neuroscience, Annual Review of Neuroscience, Behavioral and Brain Sciences, Neuron, Trends in Neurosciences, Brain, Cerebral Cortex, Neuropsychology Review, Current Opinion in Neurobiology, Journal of Neuroscience, NeuroImage). We then subsetted the MEDLINE publication metadata based on this list of selected journals.

These steps resulted in a total of 166,979 records for those 15 top journals between the year 2005-2017 which were included for further analysis.

For comparison with our publication data, we also acquired data on the percentage of NIH R01 grants in the U.S. and the percentage of MRC research grants in the U.K. awarded to women within this time period. This data was obtained from the NIH data book ("NIH Data Book," 2017) and MRC success rate data ("Medical Research Council 2016/17 Grant and Fellowship application success rates," 2018), respectively, in aggregated forms.

Gender Determination

Due to the large quantities of publication records, manually classifying author gender is infeasible. Instead, we estimated author's gender using genderizeR, the genderize.io API for R, (Wais, 2016) . The genderize.io database currently contains 216286 distinct first names and gender self-report data from social media platforms across 79 countries and 89 languages. Based on each unique first name, it outputs a predicted gender as well as a probability estimation of the prediction. For the final analysis, we included only authors with names in this database having a gender prediction with higher than 60% probability. Approximately 8% of all entries were excluded based on this criterion.

Analysis

To estimate the overall representation of women in each journal, we first calculated the overall percentage of female first and last author for each journal across the entire time range. To estimate the association between author gender ratios and journal profile, we calculated the Spearman's rank order correlation between percentage of female first and last author with journal impact factor.

To further unpack the trends of female representation over time, we also regressed the percentage of female first and last authors in each journal on time (measured in years). The

resulting slope is used as an indicator of the rate of change of female authorship in a given journal.

Results

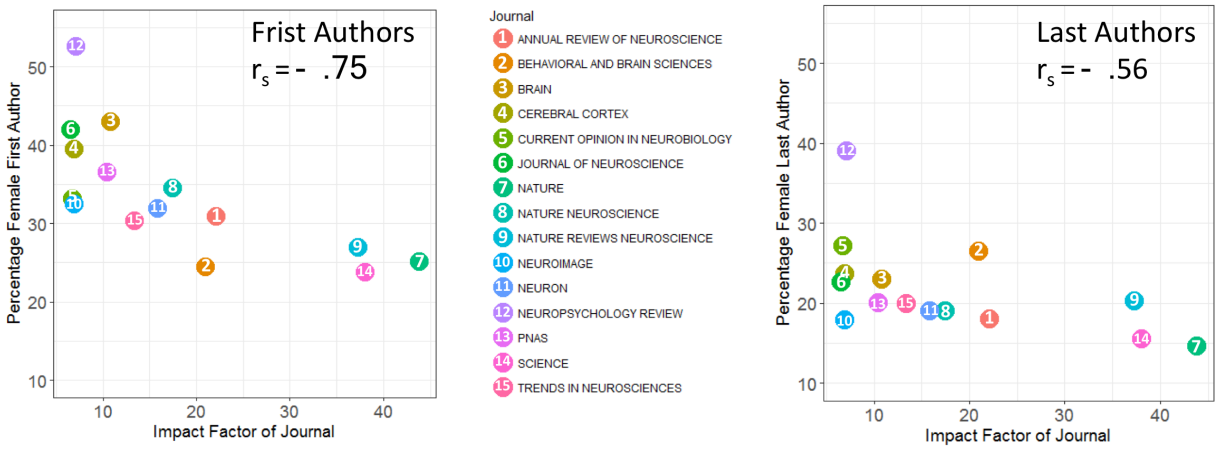

Figure 1. Percentage of female first and last authors between 2005-2017 vs. Journal's 5-year impact factor.

As shown in Figure 1, between 2005-2017, percentage of female last authors are highest in Neuropsychology Review (39.04%) and Current Opinion in Neurobiology (27.19%), and lowest in Nature (14.64%) and Science (15.53%). This pattern of results is similar for first authors, with Neuropsychology Review (52.58%) and Brain (43.01%) having the highest percentage of females, and Nature (25.22%) having the lowest. Also note that the percentage of female last authors in almost all journals (except for Neuropsychology Review) is lower than the percentage of females awarded prestigious grants such as NIH RO1 (~30%, also see grey line in Figure 2 Panel B).

The percentage of both female first and last authors displayed a strongly negative association with journal impact factor (first author $r_s = -0.75$, $p < .01$, last author $r_s = -0.56$, $p < .05$). More prestigious journals have lower representation of female first and last authors. This trend holds even when excluding the multidisciplinary journals such as Nature, Science, and PNAS (first

author r_s = -0.65, $p < .05$, last author r_s = -0.32, *ns*). Even within the field of neuroscience, a higher impact is associated with a lower female representation.

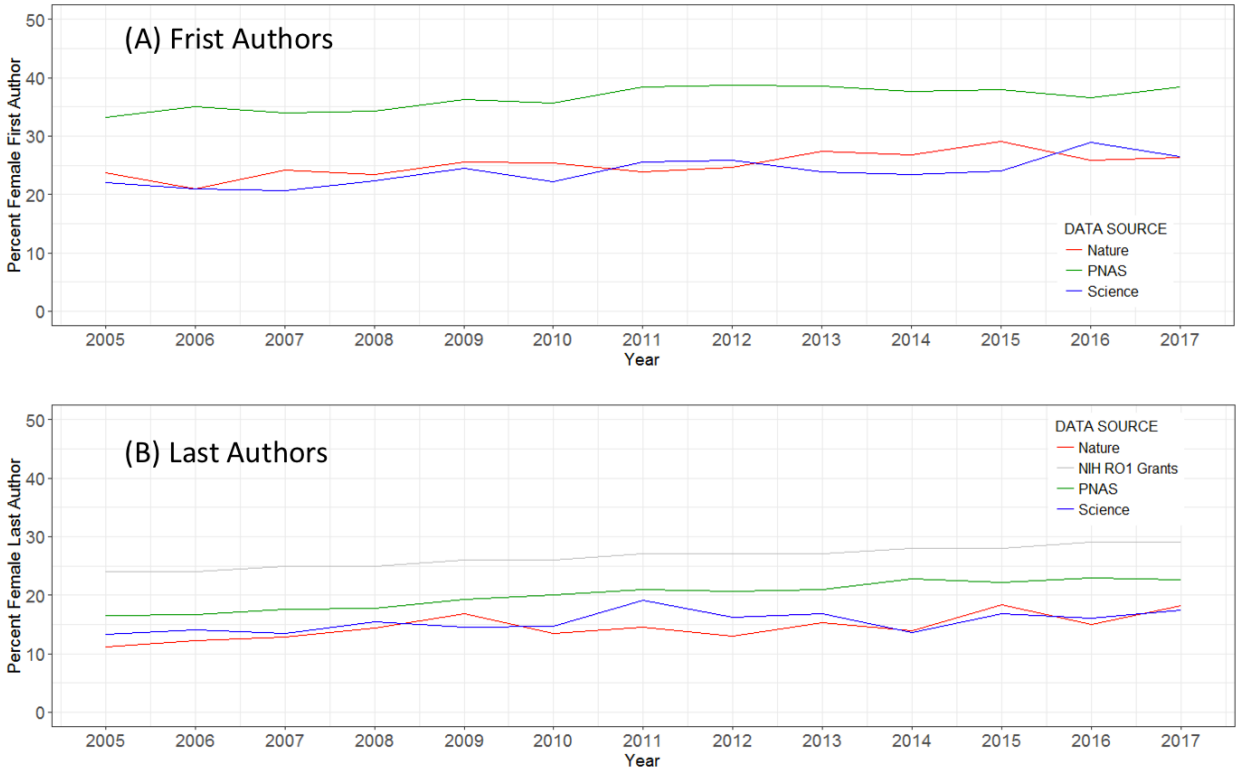

Figure 2. Change in percentage of female first (Panel A) and last (Panel B) authors overtime

As shown in Figure 2 and Table 1, the percentage of women first and last authors has increased at less than 1% per year for almost all journals (except for the journal Brain). While some journals, such as Brain, has a steady increase of female representation of over 1 percent per year, journals such as Nature Neuroscience on average shows a decrease of the percentage of female last author year by year (average decrease of -0.11% per year).

Journal	% Change in Female First Author Per Year (Slope)	% Change in Female Last Author Per Year (Slope)
Neuropsychology Review	1.85	-0.36
Nature Neuroscience	0.20	-0.11
Nature Reviews Neuroscience	0.97	0.08
Trends in Neurosciences	0.59	0.19
Neuron	0.23	0.22
Annual Review of Neuroscience	0.98	0.22
Science	0.47	0.27
Current Opinion in Neurobiology	0.22	0.31
Nature	0.40	0.40
Behavioral and Brain Sciences	1.80	0.44
Journal of Neuroscience	0.60	0.53
PNAS	0.40	0.58
NeuroImage	0.86	0.77
Cerebral Cortex	0.76	0.78
Brain	1.42	1.03

Table 1. Percentage change for female first and last authors per year.

Discussion

Like it or not, publication in high-profile journals remains an important gateway for career advancement. Consequently, the under-representation of women in high profile journals impacts thousands of talented scientists. Publishing houses have the same legal responsibility to avoid implicit (Raymond, 2013) and explicit discrimination as Microsoft, Google, and Walmart.

It is now well past time for high-impact journals to begin implementing evidence-based procedures to remove sources of bias throughout both the editorial and the reviewing process for original scientific articles. We would recommend three obvious first steps. First, all journals should collect gender and minority statistics on submitted papers and should make these data publically available. Second, journals should use mandatory double-blind reviewing. Results

from other disciplines suggest that double-blind reviewing procedures significantly increase the proportion of female lead research articles (Budden et al., 2008). Finally, reviewers should be provided with clear guidance about review criteria, as is done for NIH review panels.

References

Advisors, M. (2017). Society for Neuroscience Survey of Neuroscience Departments & Programs.

Analytics, C. (2016). 2016 Journal Citation Reports®

Budden, A. E., Tregenza, T., Aarssen, L. W., Koricheva, J., Leimu, R., & Lortie, C. J. (2008). Double-blind review favours increased representation of female authors. *Trends in Ecology & Evolution, 23*(1), 4-6. doi: 10.1016/j.tree.2007.07.008

Conley, D. (2005). Women's efforts are more than a drop in the ocean. *Nature, 438*(7071), 1078. doi: 10.1038/4381078c

Gender imbalance in science journals is still pervasive. (2017). *Nature, 541*(7638), 435-436. doi: 10.1038/541435b

Medical Research Council 2016/17 Grant and Fellowship application success rates. (2018).

MEDLINE/PubMed Data. (2017).

Nature's sexism. (2012). *Nature, 491*(7425), 495.

NIH Data Book. (2017).

Promoting diversity in neuroscience. (2018). *Nature Neuroscience, 21*(1), 1. doi: 10.1038/s41593-017-0052-6

Raymond, J. (2013). Sexist attitudes: Most of us are biased. *Nature, 495*(7439), 33-34. doi: 10.1038/495033a

Shen, H. (2013). Inequality quantified: Mind the gender gap. *Nature, 495*(7439), 22-24. doi: 10.1038/495022a

Wais, K. (2016). genderizeR: Gender Prediction Based on First Names. R package version 2.0.0.