

COVID-19 effect on the gender gap in academic publishing

Journal of Information Science

1–6

© The Author(s) 2022

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/01655515211068168

journals.sagepub.com/home/jis**Dariusz Jemielniak** 

Kozminski University, Poland

Agnieszka Sławska

Kozminski University, Poland

Maciej Wilamowski

Faculty of Economic Sciences, University of Warsaw, Poland

Abstract

The authors wanted to verify a popular belief that women scholars have been disproportionately affected by the COVID-19 pandemic. We studied the first names of authors of 266,409 articles from 2813 journals in 21 disciplines, and we found no significant differences between men and women in publication patterns between 2021, 2020, and 2019 overall. However, we found significant differences in publication patterns between gender in different disciplines. In addition, in disciplines where the proportion of women authors is higher, there are fewer single-authored articles. In the multi-author articles if the first author is female, there is more gender balance among authors, although there are still fewer women co-authors.

Keywords

bibliometrics; COVID-19; gender differences; publication counts; publication patterns

1. Introduction

There is a reasonable assumption in the popular media [1,2], as well as in academic journals and reports [3,4] that faculty parents, and especially early career women [4], have been deprived of time for research and publishing as a result of COVID-19 lockdowns.

Challenges for women academics during the pandemic, which demands remote education of children, are paramount [5]. Its effects on childcare, housework, and working from home conditions affect them disproportionately [6], adding to already preexisting biases [7].

Early data from some journals [8–11], as well as from preprints [12,13] indicate that there are significant gender disparities in publication patterns, even though some publications claim that the gender gap is not worse during the COVID-19 pandemic [14]. If it indeed is so, the situation has significant consequences for academic careers. The qualitative [15] and quantitative [16,17] studies of perceptions of scholars on this phenomenon clearly show that academics believe that women are disproportionately affected by lockdowns. However, there have not been any large data set studies of academic output which could confirm if women actually do publish less during the pandemic. In order to study the gender gap in academic publications under lockdown circumstances, we analysed 266,409 articles from the Springer-Nature database from 2813 journals in 21 disciplines (scraped with ParseHub tool, from link.springer.com with a filter set for articles).

In our study, we included only disciplines with at least 300 articles in 2021. We focused on a sample of publications from 2019, 2020, and January 2021. Our assumption was that while all publications from 2019 by default were not affected by the pandemic, the publications from January 2021 likely may have been, as an article review process takes

Corresponding author:

Dariusz Jemielniak, Kozminski University, Jagiellonska 59, Warszawa 00159, Poland.

Email: djemielniak@cyber.harvard.edu

3–6 months on average in Springer journals [18]. One possible weakness of our study is that it does not account for papers that had been written before the pandemic began, got rejected elsewhere, and later went into a review process and were successfully published in 2021. However, given the fact that the first wave of lockdowns started in March 2020, the 10-month period should be enough to expose any significant differences.

We examined author gender data by studying the first name of all named authors [19]. To identify the authors' gender, we performed the following procedure for each name from the collected database: First, we removed all names that were in a form of initials (in other words, the ones that were just one letter long). Next, we used gender guesser library to determine the gender. If a name was identified as uniquely and universally male or female, this value was used. For all the names that were not yet identified, we used gender guesser on a name where diacritic were converted to ASCII (using `unidecode` python package). Next, we looked for names that were further unidentified as males or females in a database collected from <https://github.com/MatthiasWinkelmann/name-database> in both original and simplified form. In the last step, we tried to identify gender of remaining names using database of Chinese names <https://github.com/psychbruce/ChineseNames>. In this data source, there is no direct indication if a name is used by males or females, but the number of males and females that are using this name is stated. We decided to mark an author as a male (female) when a name was given to males (females) more than 80% of the time.

One shortcoming of our approach is that in some journals, there is a practice of using initials only. In all disciplines except physics, this phenomenon affected less than 25% of the publications. In physics, publications with initials accounted for 89% of all, and thus we omit this field in our results. In the end, in our sample of author names, we identified 471,462 (37%) as male, 248,672 (20%) as female, and 544,420 (43%) were unidentified. It is important to note that we did not de-duplicate the names: in other words, individuals who published many articles, counted multiple times. We believe that this limitation does not affect the results, given the very low number of individuals who publish a large number articles in any given year.

We compared the female-to-male authors' ratio in 20 disciplines by the publication years (where data from 2021 comes mainly from January), assuming a standard difference in difference analytical method. To account for possible seasonality in the ratio of the number of female-to-male authors, we also compared its values for publications published only in January across years. We were most interested in major variations in patterns between the studied years, as it has already been shown that stable gender differences in publication numbers can be observed within disciplines [20].

2. Results

Our findings indicate that overall, there are no significant differences between men and women in publication patterns for the studied periods. The results were consistent across all four comparisons: year-to-year, January to January, 2021 to 2019, and 2021 to 2020, even though we describe the year-to-year results in this report for brevity. The distribution of the ratio of women authors is symmetrical around zero (dropped in half of the disciplines, while it increased in the others). This result indicates that there is no evidence that the share of female authors across different disciplines dropped due to the pandemic lockdown.

However, we found significant differences in publication patterns between gender in different disciplines. Moreover, the changes are stable between comparisons for 2021 to 2019 and 2021 to 2020 (see Figure 1). Out of 17 fields where we observed a year-to-year change of the female-to-male ratio greater than 5%, only in two of them, the direction of changes was different across years. When we rank the fields by the percentage change of the ratio of female to male authors for 2021/2019 and 2021/2020, the average difference of rank for a field is 2.

The highest increases of female to male authors' ratio in 2021 could be observed in geography (42.6% increase when compared to 2019 and 32% when compared to 2020), dentistry (27.7% and 19.8%, respectively), energy (25% and 7%, respectively) – see Table 1. The results for dentistry are especially interesting when the fact that medicine, in general, is a discipline with a faster publication turnaround, with an overall median submission-to-acceptance time of 123 days and acceptance-to-publication time of 68 days [21].

The highest decrease in women authors' proportion was visible in psychology (– 74.4% when compared to 2019 and – 12.3% when compared to 2020), mathematics (– 12.9% and – 17.5%, respectively) and philosophy (– 11.3% and – 10.3%, respectively). The results for philosophy and mathematics are interesting, as these disciplines do not typically require experimental studies with human subjects, and as such should be less prone to be affected by the pandemic. In addition, in the case of mathematics and psychology, we also observed a drop in the number of articles with women as first authors, as well as the number of articles with women as co-authors of articles with a man as the first author. These drops may suggest that in these fields, a culture of gender bias is more widespread than in the others.

In 2021, the highest proportion of women to men authors was visible in education (1.08), social sciences (0.8236), and dentistry (0.806), and the lowest in mathematics (0.203), psychology (0.214), and engineering (0.247).

Table 1. Changes in ratio in female to male authors, number of female to male first authors, and number of female to male authors when first author is male.

Field/year	Number of female to male authors					Number of female to male first authors					Number of female to male authors when first author male				
	2019	2020	2021	% change 19/21	% change 20/21	2019	2020	2021	% change 19/21	% change 20/21	2019	2020	2021	% change 19/21	% change 20/21
Psychology	0.84	0.24	0.21	−74.4%	−12.3%	0.74	0.24	0.20	−73.0%	−17.1%	0.33	0.12	0.10	−68.5%	−15.4%
Mathematics	0.23	0.25	0.20	−12.9%	−17.5%	0.23	0.25	0.19	−18.0%	−22.0%	0.09	0.12	0.10	7.7%	−19.7%
Philosophy	0.46	0.46	0.41	−11.3%	−10.3%	0.38	0.38	0.39	1.3%	1.3%	0.24	0.25	0.16	−35.0%	−36.5%
Environment	0.62	0.62	0.58	−7.2%	−6.6%	0.79	0.76	0.68	−13.6%	−10.6%	0.33	0.33	0.33	−2.1%	0.0%
Earth sciences	0.34	0.34	0.31	−6.3%	−6.5%	0.42	0.39	0.35	−17.3%	−12.0%	0.18	0.19	0.19	3.9%	−1.1%
Education	1.10	1.15	1.08	−2.0%	−6.2%	1.12	1.23	0.98	−12.1%	−19.8%	0.51	0.52	0.55	8.1%	5.2%
Biomedicine	0.72	0.73	0.70	−2.6%	−3.8%	0.94	0.95	1.02	9.2%	7.7%	0.45	0.44	0.45	0.7%	3.0%
Engineering	0.25	0.26	0.25	−0.8%	−5.4%	0.29	0.32	0.28	−4.2%	−13.8%	0.11	0.12	0.12	8.1%	2.6%
Statistics	0.30	0.32	0.30	−0.3%	−5.0%	0.30	0.34	0.22	−25.8%	−33.5%	0.14	0.15	0.22	53.6%	40.5%
Social sciences	0.81	0.85	0.82	1.5%	−2.6%	0.82	0.81	0.85	3.2%	5.6%	0.32	0.33	0.37	13.6%	12.6%
Life sciences	0.61	0.62	0.62	1.1%	−0.8%	0.78	0.82	0.83	6.4%	2.1%	0.35	0.36	0.34	−4.8%	−7.4%
Computer science	0.27	0.29	0.29	5.5%	−2.0%	0.32	0.41	0.33	1.9%	−19.5%	0.13	0.13	0.13	−1.5%	0.8%
Medicine	0.60	0.60	0.61	1.7%	1.8%	0.67	0.72	0.74	10.9%	4.1%	0.31	0.33	0.34	10.5%	1.5%
Economics	0.34	0.33	0.35	2.4%	6.2%	0.35	0.33	0.30	−16.0%	−9.5%	0.16	0.16	0.12	−26.2%	−23.9%
Chemistry	0.58	0.59	0.62	6.9%	4.6%	0.77	0.74	0.86	12.5%	16.0%	0.29	0.29	0.32	9.3%	10.4%
Materials sciences	0.38	0.31	0.37	−2.6%	18.6%	0.46	0.38	0.47	2.8%	25.3%	0.18	0.15	0.16	−13.3%	1.9%
Business and management	0.36	0.40	0.43	19.3%	8.5%	0.39	0.45	0.44	11.5%	−2.2%	0.17	0.18	0.20	17.1%	12.4%
Energy	0.38	0.44	0.47	25.1%	7.0%	0.44	0.54	0.54	22.8%	0.4%	0.19	0.24	0.24	27.6%	−3.3%
Dentistry	0.63	0.67	0.81	27.7%	19.8%	0.74	0.83	1.47	100.1%	77.0%	0.35	0.37	0.35	0.3%	−6.2%
Geography	0.32	0.35	0.46	42.6%	32.0%	0.32	0.36	0.55	74.6%	51.9%	0.20	0.20	0.20	0.0%	−3.0%

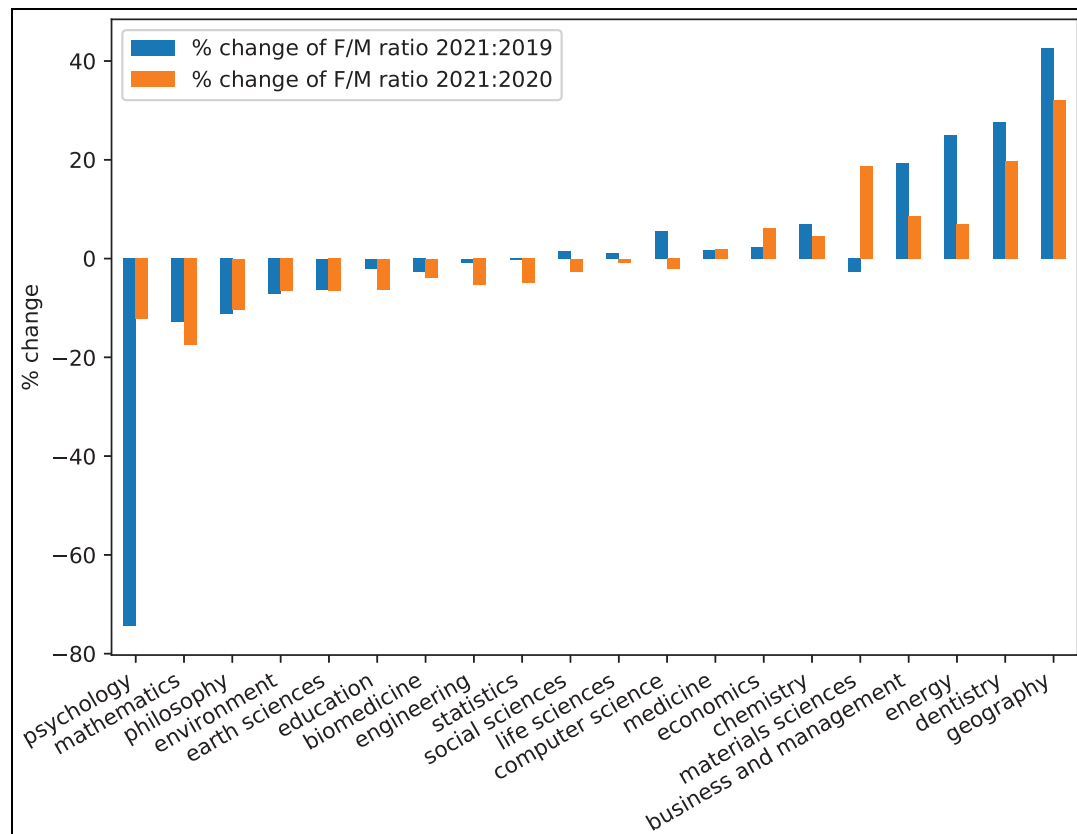


Figure 1. Percentage of women to men authors: in blue, change for 2021 to 2019, in orange change for 2021 to 2020.

The highest increase of women as first authors in 2021 was visible in dentistry (100.1% when compared to 2019 and 77% when compared to 2020) and geography (74.6% and 51.9%, respectively). The highest decrease of women as first authors was observable in statistics (−25.8% and −33.5%, respectively), and mathematics (−18% and −22%, respectively).

In 2021, the highest proportion of women to men as first authors was visible in dentistry (1.473) biomedicine (1.02) and education (0.982), and the lowest in mathematics (0.191), psychology (0.199) and statistics (0.224).

Within disciplines, we identify a strong positive correlation ($r = 0.84$) between the proportion of women authors and the proportion of women first authors while comparing results from 2019 to 2021. The identified effect is most visible in the fields of mathematics, psychology, and philosophy. On the contrary, in dentistry, the share of female authors and the share of the first female authors both increased. In most disciplines, the results are consistent with the analysis conducted for 2020–2021, and the correlation is as strong as for 2019–2020 ($r = 0.803$).

We also identify a negative correlation ($r = -0.304$ and $r = -0.562$) across disciplines between female to male authors proportion and the proportion of single-authored articles in the discipline. In other words, in disciplines where the proportion of women authors is higher, there are fewer single-authored articles.

The first author's gender appears to affect the women to men ratio among co-authors. For the articles with more than one author we calculated, per field, the ratio of the number of identified females and males among co-authors (authors not in the first position) in both cases when the first author was female and male. In the multi-author articles if the first author is female, there is more gender balance among authors – average ratio across years and fields 0.68 when female was the first author and 0.40 when the first author was male, although there are still fewer women co-authors (both ratios are lower than one).

In the three disciplines that experienced the highest decrease in overall women authors' proportion (psychology, mathematics and philosophy), the proportion of women to men co-authoring an article with a man as the first author also decreased most: in psychology (−68.5% when compared to 2019 and −15.4% when compared to 2020), mathematics (7.7% and −19.7%, respectively) and philosophy (−35% and −36.5%, respectively).

Our results do not offer a complete picture, as the effect of the pandemic lockdown should be fully visible when the whole 2021 data are available. We also note that one important factor likely significantly affecting publishing success may be parenthood: it is possible that academics who do not have young children have experienced less disruption of work than those who do, irrespective of their gender. However, we do not have access to data about the number and age of children for the studied sample.

Overall, our results clearly indicate that COVID-19 bias in gender publication patterns is not clear, the picture is complicated, and calls for further studies.


Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

ORCID iD

Dariusz Jemielniak  <https://orcid.org/0000-0002-3745-7931>

References

- [1] Kitchener C. Women academics seem to be submitting fewer papers during coronavirus. ‘Never seen anything like it’, says one editor. *The Lily*, 24 April 2020, <https://www.thelily.com/women-academics-seem-to-be-submitting-fewer-papers-during-coronavirus-never-seen-anything-like-it-says-one-editor/>
- [2] Pettit E. Covid-19 has robbed faculty parents of time for research. especially mothers. *The Chronicle of Higher Education*, 26 January 2021, <https://www.chronicle.com/article/covid-19-has-robbed-faculty-parents-of-time-for-research-especially-mothers> (accessed 1 February 2021).
- [3] Deryugina T, Shurchkov O and Stearns JE. *COVID-19 disruptions disproportionately affect female academics*. Working paper 28360. National Bureau of Economic Research. DOI: 10.3386/w28360.
- [4] Cardel MI, Dean N and Montoya-Williams D. Preventing a secondary epidemic of lost early career scientists. Effects of COVID-19 pandemic on women with children. *Ann Am Thorac Soc* 2020; 17: 1366–1370.
- [5] Gabster BP, van Daalen K, Dhatt R et al. Challenges for the female academic during the COVID-19 pandemic. *Lancet* 2020; 395: 1968–1970.
- [6] Yildirim TM and Eslen-Ziya H. The differential impact of COVID-19 on the work conditions of women and men academics during the lockdown. *Gend Work Organ*. Epub ahead of print 19 August 2020. DOI: 10.1111/gwao.12529.
- [7] Krawczyk M and Smyk M. Author’s gender affects rating of academic articles: evidence from an incentivized, deception-free laboratory experiment. *Eur Econ Rev* 2016; 90: 326–335.
- [8] McCormick CM. Disparities in the toll of the COVID-19 pandemic on publishing: evidence from submissions to *Hormones and Behavior*. *Horm Behav* 2020; 124: 104814.
- [9] Ribarovska AK, Hutchinson MR, Pittman QJ et al. Gender inequality in publishing during the COVID-19 pandemic. *Brain Behav Immun* 2021; 91: 1–3.
- [10] Pinho-Gomes A-C, Peters S, Thompson K et al. Where are the women? Gender inequalities in COVID-19 research authorship. *BMJ Glob Health* 2020; 5: e002922.
- [11] Bell ML and Fong KC. Gender differences in first and corresponding authorship in public health research submissions during the COVID-19 pandemic. *Am J Public Health* 2021; 111: 159–163.
- [12] Muric G, Lerman K and Ferrara E. COVID-19 amplifies gender disparities in research. arXiv [cs.DL], <http://arxiv.org/abs/2006.06142> (2020).
- [13] Squazzoni F, Bravo G, Grimaldo F et al. No tickets for women in the COVID-19 race? A study on manuscript submissions and reviews in 2347 Elsevier journals during the pandemic. *SSRN Electron J*. Epub ahead of print 2020. DOI: 10.2139/ssrn.3712813.
- [14] Cushman M. Gender gap in women authors is not worse during COVID-19 pandemic: results from research and practice in thrombosis and haemostasis. *Res Pract Thromb Haemost* 2020; 4: 672–673.
- [15] Górska AM, Kulicka K, Staniszevska Z et al. Deepening inequalities: what did COVID-19 reveal about the gendered nature of academic work? *Gend Work Organ* 2021; 28: 1546–1561.
- [16] Myers KR, Tham WY, Yin Y et al. Unequal effects of the COVID-19 pandemic on scientists. *Nat Hum Behav* 2020; 4: 880–883.
- [17] Gao J, Yin Y, Myers KR et al. Potentially long-lasting effects of the pandemic on scientists. *Nat Commun* 2021; 12: 6188.

- [18] SpringerNature. Timescale to publish an article for a Springer journal, <https://support.nature.com/en/support/solutions/articles/6000131708-timescale-to-publish-an-article-for-a-springer-journal> (2020, accessed 4 February 2021).
- [19] Berg J. Examining author gender data. *Science* 2019; 363: 7.
- [20] Peñas CS and Willett P. Brief communication: gender differences in publication and citation counts in librarianship and information science research. *J Inf Sci Eng* 2006; 32: 480–485.
- [21] Sebo P, Fournier JP, Ragot C et al. Factors associated with publication speed in general medical journals: a retrospective study of bibliometric data. *Scientometrics* 2019; 119: 1037–1058.