an open access journal

RESEARCH ARTICLE

Greater female first author citation advantages do not associate with reduced or reducing gender disparities in academia

Mike Thelwall and Pardeep Sud

Statistical Cybermetrics Research Group, University of Wolverhampton, UK

Keywords: academic careers, citation impact, gender bias, research evaluation

ABSTRACT

Ongoing problems attracting women into many Science, Technology, Engineering and Mathematics (STEM) subjects have many potential explanations. This article investigates whether the possible undercitation of women associates with lower proportions of, or increases in, women in a subject. It uses six million articles published in 1996–2012 across up to 331 fields in six mainly English-speaking countries: Australia, Canada, Ireland, New Zealand, the United Kingdom and the United States. The proportion of female first- and last-authored articles in each year was calculated and 4,968 regressions were run to detect first-author gender advantages in field normalized article citations. The proportion of female first authors in each field correlated highly between countries and the female first-author citation advantages derived from the regressions correlated moderately to strongly between countries, so both are relatively field specific. There was a weak tendency in the United States and New Zealand for female citation advantages to be stronger in fields with fewer women, after excluding small fields, but there was no other association evidence. There was no evidence of female citation advantages or disadvantages to be a cause or effect of changes in the proportions of women in a field for any country. Inappropriate uses of career-level citations are a likelier source of gender inequities.

Citation: Thelwall, M., & Sud, P. (2020). Greater female first author citation advantages do not associate with reduced or reducing gender disparities in academia. *Quantitative Science Studies*, *1*(3), 1283–1297. https://doi.org/10.1162/qss_a_00069

DOI: https://doi.org/10.1162/qss_a_00069

Received: 20 February 2020
Accepted: 04 April 2020

Corresponding Author:
Mike Thelwall
m.thelwall@wlv.ac.uk

Handling Editor:
Ludo Waltman

1. INTRODUCTION

Female authors are underrepresented in almost all countries and most institutions (Guglielmi, 2019), but not all broad fields. Science, Technology, Engineering and Mathematics (STEM) subjects tend to have a minority of female academics (NCES, 2019a) and female students in the United States (NCES, 2019b) and probably in most Western nations (e.g., European Commission, 2019). As women tend to dominate the health care, elementary education, and the domestic sphere (HEED) broad fields (Tellhed, Bäckström, & Björklund, 2017), their STEM scarcity raises the suspicion that women are being discouraged from some areas. It is therefore important to understand the causes of female underrepresentation in STEM and academia overall. This article focuses on one hypothesized cause: the perceived value of female contributions to research, as reflected in citations to their work. Bias against women translating into undercitation for their articles has been argued to be a factor in female underrepresentation (Larivière, Ni, et al., 2013). Fewer citations for female-authored research can damage careers when citations are considered within hiring, promotion, and funding decisions. This article uses a new approach to investigate whether gendered citation differentials may have

The MIT Press

influenced the rate at which academic specialisms attract women, with a focus on six large mainly English-speaking countries.

One early speculation about the cause of female STEM underrepresentation was that women might be less capable in numeric subjects (*Guardian*, 2005) but girls slightly outperform boys in mathematics overall (Lindberg, Hyde, et al., 2010), so explanations have been sought elsewhere. Other possible causes include societal gender role expectations or personal choice. The choice factor may apply due to liking a subject (Su, Rounds, & Armstrong, 2009) or because it supports a person's wider life goals (Diekman, Steinberg, et al., 2017). Both of these factors may also influence female career preferences and so may be influential at many life stages. The rest of this section focuses on gender bias within academia.

A major review of evidence for gender bias being a substantial cause of female underrepresentation in STEM compared to other subjects did not find it credible overall (Ceci & Williams, 2011). Although some investigations have found evidence of antiwoman bias in some academic contexts, such as journal article refereeing or funding applications (Budden, Tregenza, et al., 2008), many others have found no differences (Grant, Burden, & Breen, 1997; Whittaker, 2008). Nevertheless, the existence of particularly low rates of women in engineering and computer science raises the possibility of remaining pockets of (intentional or unintentional) bias, or "chilly climates" that may alienate or disadvantage women (Stockard, Greene, et al., 2018). There is a lack of strong recent evidence for this being a widespread phenomenon, however (e.g., Jorstad, Starobin, et al., 2017). There may also be biases in aspects of academia that have not been measured and there may be a cumulative effect of small elements of bias throughout some academic subjects that are difficult to measure individually.

Bias has been previously examined though citation analysis. As people cite work that may have influenced them, it could be expected that bias against women might translate into fewer citations for their articles. Evidence in support of this has been found in a degree of gender homophily in citation practices (Mitchell, Lange, & Brus, 2013; Potthoff & Zimmermann, 2017), but it is impossible to be sure that this is not due to gender differences in methods and topic choices, even within narrow subjects (Grant, Ward, & Rong, 1987; Thelwall, Bailey, et al., 2019a, 2019b). Overall biases have also been sought by comparing the citation rates of female- and male-authored journal articles, with the assumption that a lower rate of citation for female-authored work might reflect bias against women (Larivière et al., 2013). Although this study found fewer citations per paper for female-authored research overall, subsequent studies with a normalizing technique that did not allow individual highly cited articles to dominate found the opposite overall (Thelwall, 2018). For English-speaking nations, a female citation advantage has been found for most broad fields (Thelwall, 2020a), and for most narrow fields (Thelwall, 2020c). There have also been studies of gender differences in citation rates within individual fields (e.g., Østby, Strand, et al., 2013). A detailed study of medicine found that gender differences in first and last author citation rates could be due to collaboration, journal choice, and self-citations (Andersen, Schneider, et al., 2019).

The current article investigates gender biases in citations by generating a new type of large science-wide evidence: a comparison of gendered citation advantages with gendered participation rates at the narrow field level. If bias is a major cause of female underrepresentation in some subjects, then female-authored articles would attract the fewest citations relative to those of men in areas in which they were the smallest minority. This article therefore assesses whether the proportion of female first-authored papers in a narrow subject correlates with the rate at which female first-authored papers are cited, relative to those of men. The analysis focuses on the United States, the main research producer according to the major citation indexes, and five comparable countries in terms of a (mostly) common language, similar level of economic development, and partly shared cultural heritage. Last-author genders are modeled but not analyzed separately because the role of the last author varies between fields, so it is not useful to run a cross-field analysis of this position. The five research questions address this and related issues that help assess the coherence of the main goal.

- **RQ1**: Do large English-speaking countries have similar proportions of female first authors across narrow fields? Field topics and disciplinary cultures tend to be international to some extent, so it is plausible that subjects would have similar gender proportions internationally, especially in countries with similar cultures and a shared language. Previous research has found this to be broadly true at the discipline level, but with some major exceptions (e.g., veterinary research is the most female-friendly broad field in the United States, but the least in India: Thelwall et al., 2019a, 2019b).
- **RQ2**: Are gender differences in first-author citation advantages similar across narrow fields in large English-speaking countries? As above, because disciplinary cultures are international to some extent and gender citation advantages are independent of national policies, a moderate or high positive correlation should be expected between the citation advantages of different countries with similar cultures.
- **RQ3**: Are female first-author citation advantages higher in narrow fields with a greater proportion of women? There are multiple reasons for hypothesizing that gender majority authors would have a citation advantage in a field. This is a logical outcome of gender bias (the main factor investigated here) or any degree of gender homophily in citations. More broadly, (a) assuming a normally distributed degree of interest in a subject for each gender, but with a different mean, the average level of interest would be higher for the majority gender (because this gender would include more of the tail of its distribution). If this level of interest (b) translates into greater effort in the subject and (c) this greater effort translates into more citations, then research from the majority gender would tend to be more cited. Conversely, minority genders might attract more citations if there is a degree of awareness from the majority that the minority need to be nurtured for the health of the discipline.
- **RQ4**: Are female first author citation advantages higher in narrow fields with a greater *increase* in the proportion of women? If citation advantages influence female careers, there should be a tendency for the proportion of women to increase most in fields with the greatest female citation advantage.
- **RQ5**: Do the answers to RQ2, RQ3, and RQ4 depend on whether team size is ignored when calculating female citation advantages? The citation advantage questions can be addressed with either of two reasonable assumptions: that the success of a team-authored paper (which is likely to be more cited: Persson, Glänzel, & Danell, 2004) should be modeled as partly due to the work of the first author and partly due to the amount of help that he or she has received (team size); or that the success of a team-authored paper should be attributed only to the first author, who is therefore implicitly credited with putting together the team and their contributions to the success of the paper.

2. METHODS

This article is a follow up to two previous papers using the same data set. The first assessed gender differences in citation impact for seven large English-speaking nations overall 1996–2018 and in terms of citation distributions, finding a female citation advantage in six. This was mainly due to more female first-authored highly cited articles (Thelwall, 2020c). The second investigated 27 broad fields (Scopus categories) for six large English-speaking nations, 1996–2014, showing considerable variation between fields in the nature and magnitude of the citation advantages, with smaller variations between countries (Thelwall, 2020a). The current paper advances from these papers by using regression to estimate gender factors, taking into account team size, and comparing the results with the proportions of authors in narrow fields to study the relationship between gender compositions of narrow fields and gender differences in citation impact.

Scopus assigns journals to one or more narrow fields from a set of 334 within 27 broad fields (https://www.elsevier.com/solutions/scopus/how-scopus-works/content). For example, *Literature and Literary Theory* is a narrow field within the broad field *Arts and Humanities*. In the remainder of the article the term *field* refers to Scopus narrow fields, unless specified. The focus is on narrow fields rather than broad fields to give fine-grained information about gender differences, given that related fields may have different gender balances and dynamics. Analyzing broad fields (but field normalizing citations at the narrow field level) would make the analysis methods more statistically powerful if effect sizes were similar across constituent narrow fields, but the sample sizes are sufficient to give results for most narrow fields in most countries, so the narrow field approach seems reasonable.

2.1. Data Collection and Gender Detection

Journal article records from Scopus 1996–2012 were downloaded in November–December 2018, including author names and countries, as well as the Scopus subject areas of each article and its citation count. This data set is recycled from multiple previous studies. Scopus was used in preference to the Web of Science for its wider coverage of different subject areas (Mongeon & Paul-Hus, 2016). The starting year is the first after a Scopus expansion and so is a logical choice. The final year, 2012, allows at least 5 years for each article to attract citations. Although 3 years is enough in most cases (Abramo, Cicero, & D'Angelo, 2011), 2 extra years were included to allow some of the slower maturing areas of the social sciences to be included. For each country, articles were discarded unless all affiliations in Scopus were from that country. Although international collaboration is important, it influences citation rates (Didegah & Thelwall, 2013) and so was excluded to avoid confounding effects.

As before (Thelwall, 2020c), author genders were estimated from their first names, as registered in Scopus, using a list of gendered first names from the US census 2010 and results from gender-api.com. This method has an accuracy of 98.5% based on a manually classified sample of U.S. academics from 2017 collected from a previous paper (Thelwall et al., 2019b). As Canada has a substantial French-speaking minority, a similar French first name list was compiled using gender-api.com and combined with the English list, deleting conflicting names. French first names were used only when gender-api.com had at least 100 records and reported at least 90% with the same gender.

Only first and last author genders were considered for the regressions to estimate gender citation advantages because the first author is usually the main contributor in all fields, although the last author can make a substantial secondary contribution in some (Larivière, Desrochers, et al., 2016). For consistency and due to the lack of systematic information about which Scopus fields contain articles in which the last author usually plays a substantial role, last authors were modeled for all fields. In fields where the last author place is irrelevant, modeling their gender adds a spurious variable that should not influence the results substantially. Although alphabetical ordering probably occurs to some extent in most fields, such as mathematics, economics, and large physics experiments (Levitt & Thelwall, 2013; Waltman, 2012), this is irrelevant for solo research and monogender teams and serves overall to add noise to the results for these fields, weakening rather than biasing the results. The overall rate of gender detection (i.e., the percentage of papers with both first and last author genders detected) was Australia: 57.1%, Canada: 55.0%, Ireland: 55.2%, New Zealand: 56.1%, United Kingdom: 55.4%, and United States: 57.5%.

2.2. Regressions to Estimate Gendered Citation Advantage Indicators

Female citation advantages were estimated using linear regression applied to models where the (narrow) field normalized citation count is the dependent variable and the independent variables are first author gender, last author gender, team size (1, 2, 3, 4, 5+ authors), and publication year. Field and year normalization are necessary for citation counts because they naturally vary by year and field (Waltman, van Eck, et al., 2011).

Various publication-related variables are known to associate with citation counts, but were not included in the regression model because they are under the control of the author. These include journal impact factor, title length, and number of references (Tahamtan, Afshar, & Ahamdzadeh, 2016). For example, if women are more cited primarily because they publish in higher impact journals, construct longer titles or add more references, then gender citation differences would be hidden by including these variables in the regressions. The number of authors is included because additional authors after the first contribute to a paper in a way that cannot be directly attributed to the first author. Similarly, the gender of the last author is included because in some fields they make a substantial contribution. It is possible that gender differences in this role might influence citation impact in these teams in a way that it is not independent of first-author gender (e.g., senior male last authors in medicine often leading teams with junior female first authors).

The field normalized citation count used is the normalized log-transformed citation score (NLCS), which is the log-transformed citation count $ln(1 + c)$ for an article divided by the average of all log-transformed citation counts $ln(1 + c)$ for all articles in the same narrow field and year (Thelwall, 2017). If an article is in multiple narrow fields then its logged citation count is instead divided by the average of all the field average logged citation counts. The log transformation prevents individual highly cited articles from dominating the results. It is slightly better than the alternative of using negative binomial regression (Thelwall, 2016). The use of NLCS instead of a count data model also allows the method to normalize simultaneously for multiple fields so that interdisciplinary research is fairly and naturally treated.

Publication year is included as a variable because the average impact of each country varies over time relative to the world average and including publication year allows the gender variables to ignore this factor. Multilevel modeling could also solve this problem, but there are sufficient data to avoid having to make the extra assumptions required for this approach.

Team size was modeled separately for 1, 2, 3, 4, and 5+ authors rather than by using a log or other formula because there is no accepted relationship. As the regressions were fitted separately for each country and year (a theoretical maximum of 334 × 6 regressions, one for each of the 334 Scopus narrow fields and each of the six countries, although the valid maximum for any country was 331), this also allowed the relationship between team size and normalized citations to vary between fields and countries.

For consistency, the same set of team size variables was used for all fields. This is a problem for fields in which team sizes larger than five have substantial differences in average citation impact. For example, perhaps huge team physics research (e.g., >1,000 authors) has much higher citation impact than teams of 10–100 authors due to more expensive equipment, or large medical teams (e.g., 25+ authors) tend to be well-funded international consortia producing high-impact work. As there does not seem to be a way to deal with this issue effectively (Mongeon, Smith, et al., 2017; Thelwall, 2020b), this is a methodological limitation.

The linear regression fitted to each country and year combination was therefore compiled from binary independent variables as follows, where $F = 1$ if the first author is female (otherwise 0), $A_i = 1$ if the article has i authors (or $\geq i$ if $i = 5$), $L = 1$ if the last author is female, and $P_i = 1$ if the article was published in the year i (if all P_i are 0 then the article is from 1996).

$$\mathrm{NLCS} = \alpha + \beta_1 \mathrm{F} + \beta_2 \mathrm{A}_2 + \beta_3 \mathrm{A}_3 + \beta_4 \mathrm{A}_4 + \beta_5 \mathrm{A}_{5+} + \beta_6 \mathrm{L} + \beta_7 \mathrm{P}_{1997} + \ldots + \beta_{22} \mathrm{P}_{2012}$$

A reduced regression model was also fitted to each field/year combination, without the author numbers variables to help address the last research question.

$$\mathrm{NLCS} = \alpha + \beta_1 \mathrm{F} + \beta_2 \mathrm{L} + \beta_3 \mathrm{P}_{1997} + \ldots + \beta_{18} \mathrm{P}_{2012}$$

Again, as before, the linear regression model fitted was a weighted least squares variant of ordinary least squares regression (Yohai, 1987), in R (Rousseeuw, Croux, et al., 2019), which can cope with substantial variance heterogeneity in publication year and (usually) minor but sometimes statistically significant variance heterogeneity in the other variables. Field/country combinations where the regression did not converge due to a lack of data were discarded.

The female regression coefficient in each of the models is the estimated average citation advantage of women in the field and country 1996–2012. For example, a β_1 value of 0.001 indicates that female first-authored research tended to receive a 0.1% higher NLCS (i.e., a 0.1% higher ratio of log-transformed citations to the world average). Similarly, negative values indicate a male citation advantage. Some of these coefficients were not statistically significant, but all were included in the data irrespective of significance level because they are estimates of the female citation advantage or disadvantage, whether statistically significant or not.

In summary, the gender indicators used for the current paper were estimates of the proportion of female authors in each field and six countries, together with two regression-derived estimates of the female first-author citation advantage, with and without team size as a contributory factor. The analyses were conducted twice: once for all the data and once for fields with at least 50 gendered articles in 1996 and at least 50 gendered articles in 2012, so that the number of articles analyzed is larger and the proportion gender change between 1996 and 2012 is more accurate. The full models were less likely to converge than the reduced models due to the additional independent variables, so there are more results for the reduced models (Table 1).

2.3. Analysis

The main analysis is an investigation of the linear regression results and gender proportions using correlation tests. Spearman correlations were used instead of Pearson correlations because the variables were not normally distributed. Bonferroni correction procedures (Hochberg, 1988) for multiple simultaneous tests are not reported because the tests could be interpreted separately for individual countries or collectively for countries as a set, which

Table 1. Summary of the amount of data analyzed for each country. The last three columns are for field/country combinations with at least 50 articles in 1996 and at least 50 in 2012 (Total = 4,968 successful regressions)

Country	Articles	Fields	Full model	Reduced model	Articles 50	Fields 50	Full model 50	Reduced model 50
Australia	374,052	330	308	310	236,595	91	90	91
Canada	509,478	328	310	313	401,210	148	148	148
Ireland	42,872	321	222	250	0	0	0	0
New Zealand	64,818	324	239	259	9,763	6	6	6
United Kingdom	937,390	328	319	321	839,293	199	199	199
United States	4,553,110	331	325	324	4,488,242	291	290	291
Total	**6,481,720**	–	–	–	**5,975,103**	–	–	–

influences the nature of any correction needed. Instead, the issue of multiple tests is discussed when relevant. Correlation tests are used here to detect science-wide patterns or tendencies. Although each field and nation will have individual characteristics, strong enough science-wide effects would translate into positive correlations.

RQ1: Do large English-speaking countries have similar proportions of female first authors across fields? This was assessed by calculating the Spearman correlation between the female first author proportions for each field between all pairs of countries. The correlations were conducted pairwise, excluding countries for which no gendered publications were found. To guard against small numbers effects, the correlations were repeated after excluding field/country combinations with fewer than 50 gendered papers in 1996 and 2012.

RQ2: Are gender differences in first author citation advantages similar across fields in large English-speaking countries? This was assessed as above but using the female first-author citation advantage instead of the gender proportion.

RQ3: Are female first-author citation advantages higher in fields with a greater proportion of women? This was assessed within each country by correlating the female first-author proportion (1996, 2012, or overall) with the female first-author citation advantage.

RQ4: Are female first-author citation advantages higher in fields with a greater increase in the proportion of women? This was assessed within each country by correlating the increase in female first-author proportions between 1996 and 2012 with the female first-author citation advantage.

RQ5: Do the answers to RQ2, RQ3, and RQ4 depend on whether team size is ignored when calculating female citation advantages? This was assessed directly for RQ2 and by evaluating RQ3 and RQ4 twice, with female first-author citation advantage estimates from each of the regression models with and without the team size independent variables.

3. RESULTS

Full results and details of regression fitting are in the supplementary materials on FigShare (about 20,000 pages at https://doi.org/10.6084/m9.figshare.9036884), together with additional analyses not included here.

Table 2. The percentage of female first-author citation advantages out of all fields where the regression converged and there was a first-author gender difference

Country	Fields with a female first-author citation advantage (full model 1)	Fields with a female first-author citation advantage (reduced model 2)
Australia	162/308 (53%)	203/310 (65%)
Canada	166/310 (54%)	203/313 (65%)
Ireland	115/222 (52%)	144/250 (58%)
New Zealand	143/239 (60%)	175/259 (68%)
United Kingdom	168/319 (53%)	240/321 (75%)
United States	200/325 (62%)	248/324 (77%)

As background information, in all countries, whichever model was used, there was a female first-author citation advantage in more fields than the reverse (Table 2), as previously found with the same data (Thelwall, 2020a).

3.1. RQ1: Do Large English-Speaking Countries Have Similar Proportions of Female First Authors across Fields?

There is a high correlation between the proportion of women in each narrow field between all 15 pairs of countries (Tables 3 and 4). Thus, a field with a relatively high proportion of women in one country would also tend to have a relatively high proportion of women in the other five. This suggests that there is a strong field-based component in gender differences that transcends nation, at least among countries with similar cultures. The lowest correlations involve the countries with the fewest articles and are probably due to small sample size issues rather than more systematic differences with the larger countries.

3.2. RQ2/5: Are Gender Differences in First-Author Citation Advantages Similar for Fields in Large English-Speaking Countries?

A tendency for positive correlations between first-author citation advantages in similar countries (Table 5) and many statistically significant small positive correlations (eight out of 15 in

Table 3. Spearman correlations for the proportion of female first authors in each Scopus narrow field 1996–2012 between pairs of large English-speaking countries. Pairwise sample sizes (between 319 and 331) are in the supplementary materials (https://doi.org/10.6084/m9.figshare.9036884)

	Canada	Ireland	New Zealand	United Kingdom	United States
Australia	.922***	.773***	.786***	.920***	.938***
Canada	1	.750***	.798***	.925***	.934***
Ireland		1	.716***	.787***	.771***
New Zealand			1	.770***	.769***
United Kingdom				1	.930***

*$p < .05$, **$p < .01$, ***$p < .001$.

Table 4. Spearman correlations for the proportion of female first authors in each Scopus narrow field 1996–2012 between pairs of large English-speaking countries. Pairwise sample sizes (between 0 and 291) are in the supplementary materials (https://doi.org/10.6084/m9.figshare.9036884). Fields are included only if they have at least 50 gendered articles in both 1996 and 2012

	Canada	Ireland	New Zealand	United Kingdom	United States
Australia	.971***	–	1.000***	.970***	.960***
Canada	1	–	.750	.972***	.976***
Ireland		–	–	–	–
New Zealand			1	.929**	1.000***
United Kingdom				1	.965***

*p < .05, **p < .01, ***p < .001.

Table 5) for the full data set gives evidence that there is a tendency for female first-author citation advantages to be to a small degree similar in the same field between similar countries. The correlations are stronger for the more robust data set with larger sample sizes, with statistically significant moderate and strong positive correlations (four out of 15 in Table 6). Assuming that the higher correlations in the latter case are due to the greater amount of data generating more accurate female first-author citation advantage estimates, this suggests that that there tends to be a moderately strong relationship between female first-author citation advantages in countries with similar cultures. This gives evidence that female first-author citation advantages are field-specific to a moderate extent.

If author numbers are left out of the regression model then there are small changes in the correlations, but the overall conclusions are not affected (Tables 7 and 8).

3.3. RQ3/4/5: Are Female First-Author Citation Advantages Higher in Fields with a Greater Proportion of, or Increase in, Women?

RQ3/5: When all available data are used, there is a small tendency for female citation advantages to correlate negatively with the proportion of women in a field (Table 9, columns 2 and 3). This correlation is statistically significant for two out of six countries, irrespective of whether the full or reduced regression models are used. The results would remain significant if a

Table 5. Spearman correlations for the female first-author citation advantages in each Scopus narrow field between pairs of large English-speaking countries. Pairwise sample sizes (between 202 and 325) are in the supplementary materials (full regression model including author counts)

	Canada	Ireland	New Zealand	United Kingdom	United States
Australia	.120*	.033	.082	.229***	.257***
Canada	1	.128	.115	.175**	.216***
Ireland		1	−.106	.190**	.142*
New Zealand			1	.052	−.023
United Kingdom				1	.270***

*p < .05, **p < .01, ***p < .001.

Table 6. Spearman correlations for the female first-author citation advantages in each Scopus narrow field between pairs of large English-speaking countries. Pairwise sample sizes (between 0 and 290) are in the supplementary materials (full regression model including author counts). Fields are included only if they have at least 50 gendered articles in both 1996 and 2012

	Canada	Ireland	New Zealand	United Kingdom	United States
Australia	.101	–	−.029	.445***	.327**
Canada	1	–	−.257	.243**	.156
Ireland		–	–	–	–
New Zealand			1	.771	−.543
United Kingdom				1	.387***

*$p < .05$, **$p < .01$, ***$p < .001$.

Table 7. Spearman correlations for the female first-author citation advantages in each Scopus narrow field between pairs of large English-speaking countries. Pairwise sample sizes (between 231 and 327) are in the supplementary materials (reduced regression model excluding author counts)

	Canada	Ireland	New Zealand	United Kingdom	United States
Australia	.167**	.015	.105	.272***	.300***
Canada	1	.097	.113	.250***	.195***
Ireland		1	.032	.181**	.082
New Zealand			1	.087	.025
United Kingdom				1	.232***

*$p < .05$, **$p < .01$, ***$p < .001$.

Table 8. Spearman correlations for the female first-author citation advantages in each Scopus narrow field between pairs of large English-speaking countries. Pairwise sample sizes (between 0 and 291) are in the supplementary materials (reduced regression model excluding author counts). Fields are included only if they have at least 50 gendered articles in both 1996 and 2012

	Canada	Ireland	New Zealand	United Kingdom	United States
Australia	.309**	–	.086	.529***	.499***
Canada	1	–	.371	.303***	.219**
Ireland		–	–	–	–
New Zealand			1	.771	.257
United Kingdom				1	.402***

*$p < .05$, **$p < .01$, ***$p < .001$.

Table 9. Spearman correlations between the proportion of female first authors 1996–2012 or the proportion change from 1996 to 2012 and the female citation advantage (full regression model 1 or reduced regression model 2). Pairwise sample sizes (between 168 and 325) are in the supplementary materials

Country	Female proportion against female citation advantage (1)	Female proportion against female citation advantage (2)	Female change against female citation advantage (1)	Female change against female citation advantage (2)
Australia	.025	.013	.057	.068
Canada	−.052	−.058	.019	.027
Ireland	−.068	.005	.031	.060
New Zealand	−.197**	−.164**	.058	.019
United Kingdom	.010	−.048	.143*	.098
United States	−.190***	−.156**	.053	.014

*$p < .05$, **$p < .01$, ***$p < .001$.

Bonferroni correction were to be applied. When only fields with at least 50 gendered articles in both 1996 and 2012 are included, the significant correlations disappear, however, as does the tendency for the correlations to be negative (Table 10). As spurious significant correlations should not be the result of noisier data (fields with fewer articles), the overall conclusion is that there is some evidence for a weak negative correlation between the proportion of women in fields that are not small and the female citation advantage (i.e., the more women, the smaller

Table 10. Spearman correlations between the proportion of female first authors 1996–2012 or the proportion change from 1996 to 2012 and the female citation advantage (full regression model 1 or reduced regression model 2). Pairwise sample sizes (between 0 and 291) are in the supplementary materials. Fields are included only if they have at least 50 gendered articles in both 1996 and 2012

Country	Female proportion – female citation advantage (1)	Female proportion – female citation advantage (2)	Female change – female citation advantage (1)	Female change – female citation xadvantage (2)
Australia	.119	.102	.158	.081
Canada	.055	.071	.022	−.002
Ireland	–	–	–	–
New Zealand	−.029	−.029	.829*	.829*
United Kingdom	.052	−.010	.079	.081
United States	−.113	−.071	.106	.050

*$p < .05$, **$p < .01$, ***$p < 0.001$.

the female citation advantage). The results for fields that are not small are consistent with there being (a) a negative long-term relationship for all six English-speaking countries, but only two are statistically significant, or (b) a negative long-term relationship in some but not others of the set analyzed. For all fields, irrespective of size, there is insufficient evidence to claim a relationship between female citation advantages and the proportion of women in a field. Thus, small fields may have a different relationship to other fields.

RQ4/5: When all available data are used, there is no statistical evidence that a female first-author citation advantage associates with an increase in the proportion of women, although the correlations are small and positive in all countries (Table 9, column 4). When fields are required to have at least 50 gendered articles in 1996 and 2012 so that the gender change calculation is more reliable (Table 10, column 5), only one correlation is statistically significant, and this would disappear if a Bonferroni correction were applied ($p = .034$). Thus, overall, the results do not give evidence of a connection between female citation advantages and increases in the proportions of women.

4. DISCUSSION

In parallel with the many possible causes of female underrepresentation in STEM subjects, there are many possible causes of gender citation differentials, other than bias. Multidisciplinary research can cause differentials if one gender is more likely to conduct multidisciplinary investigations and if the field normalization process above (simple averaging) systematically overestimates (or underestimates) the citation impact of such work. One gender may also be more likely to conduct a type of method (e.g., qualitative: Ashmos Plowman & Smith, 2011) that modifies the expected citation rate. Antibias is another possibility, if editors and referees are more generous to minority genders in a field so that their *published* work would be, on average, lower quality. Capability or level of interest may also be factors. If one gender is more capable in a subject by the time it is employed in a research capacity (e.g., by taking more after-class voluntary courses) or interested in a topic (e.g., because it supports their communal goals: Diekman et al., 2017), on average, by the time they become researchers then that gender might tend to produce better work as a result. It is also possible that multiple factors apply, partly cancelling each other out, and working differently between fields and countries. National discipline-wide gender action initiatives may also have impacted on fields. Thus, the citation correlation evidence here cannot directly prove any cause-and-effect relationship.

RQ1: The finding that the proportion of female researchers within individual narrow fields is consistent between the six countries examined is unsurprising. This issue has been reported before, although at a much broader level of granularity. For example, Table 7.3 of She Figures 2018 (European Commission, 2019) reports the proportion of female authors in six broad areas 2013–17, giving figures to one decimal place that show a high correlation between countries. This is probably a culture-specific high correlation, however, because there are known examples of countries with different gender ratios for some fields, such as the female domination of computer science in Malaysia (Othman & Latih, 2006) and the relative lack of women in veterinary science in India (Thelwall et al., 2019b).

RQ2: No previous study seems to have investigated whether there are international similarities in gender citation advantages at the broad or narrow field level, so the positive answer to the second research question is unique. There are at least two contrasting explanations of the positive answer, however. If the female citation advantage is an artifact of the Scopus categorization scheme (e.g., by mixing higher citation female-oriented specialties with lower citation male-oriented specialties within a narrow field) then this would give internationally consistent female citation advantages, at least for publications in relatively international journals. The somewhat consistent female citation advantage between fields could also be due to a subject-based tendency for women to attract more citations in specific fields, for example due to focusing on a higher citation topic, or being more capable (for whatever reason); alternatively, there may be bias in the (fewer) fields in which they are less cited.

RQ3: No previous study has investigated whether gendered citation advantages associate with higher proportions of women, so a small tendency for the opposite to occur is a new finding. This finding undermines the hypothesis in previous studies that citation bias against women may help to explain their underrepresentation in some fields. This is not evidence that citation bias does not happen or that it never affects field participation rates, since other factors affect citation rates. Nevertheless, it seems likely that citation biases, if and when they occur, are not substantial enough to affect careers sufficiently to be relevant as determinants of differing female participation rates in narrow fields. One way in which citation biases can occur is through gender homophily in citations (Mitchell et al., 2013; Potthoff & Zimmermann, 2017), but this does not seem to have a strong effect overall, given that women tend to be more cited in fields with a higher proportion of men.

RQ4: No previous study has investigated whether gendered citation advantages associate with increasing proportions of women, so the lack of evidence that this occurs is new information. This corroborates the RQ3 answer, further undermining the citation bias explanation for female underrepresentation in some fields.

RQ5: The answers to RQ2, RQ3, and RQ4 do not depend on whether it is assumed that the first author is responsible for putting together the research team, rather than modeling citation impact as partly contributed to by the size of the authorship team.

5. CONCLUSIONS

The results are broadly consistent with the absence of a systematic gender bias in citations and with this bias, if it exists, not being influential on the proportion of women in a field, at least at the level of citations per paper. Where women are in a smaller minority, their work does not seem to be less cited, and lower citation rates for women do not seem to lead to decreasing numbers of women in a field. Although this is a negative finding, it contributes to current efforts to reduce gender disparities by ruling out one possibility: Energy should not be spent on looking for gender citing biases through gender homophily in citing and it should not be assumed that the papers of female authors will be overlooked when they are a minority within a field. Other causes and solutions for disparities should be pursued instead.

Despite the main findings, at the *career* citations level (total citations, *h*-index), women still face citation-related disadvantages. This is because the greater contributions of women to society outside of formal work environments (e.g., greater childcare and other carer responsibilities leading to career gaps and period of part-time working) may well have a pernicious career effect due to inappropriate uses of bibliometrics in ranking lists and for promotions, funding, and tenure. Career-based or productivity-based indicators favor men if they do not consider the time (Van Den Besselaar, & Sandström, 2016) and resources (Ceci & Williams, 2011) available because of career decisions, such as periods of part-time working or career breaks for carer responsibilities or teaching-related duties (see also Kretschmer & Kretschmer, 2013). For example, if citations were used in recruitment and tenure but the assessors did not carefully compensate candidates for part-time working and career breaks then women would be disadvantaged, on average. Similarly, there seems still to be a passive acceptance of male-dominated citation ranking lists, such as the Google Scholar *h*-index ranking (Carter, Smith, & Osteen, 2017), perhaps Clarivate's Highly Cited Researchers list, and the male domination of prizes in many fields, including the scientific Nobel Prizes, that do not seem to take into account the abovementioned gender differences in working lives. As a practical step, all researchers should treat awards and ranking lists with great suspicion as to whether they underrepresent women (or other disadvantaged groups in society) in comparison to their proportion in the relevant scientific field.

Finally, taking into account the main findings and the above discussion of career citations, the main gender-related focus of scientometricians should be ensuring that career-level citation impact data, when used, is not biased against women.

AUTHOR CONTRIBUTIONS

Mike Thelwall: Conceptualization, Investigation, Methodology, Visualization, Writing—original draft, Writing—review & editing. Pardeep Sud: Writing—original draft, Writing—review & editing.

COMPETING INTERESTS

The authors have no competing interests.

FUNDING INFORMATION

This research was not funded.

DATA AVAILABILITY

The processed data used to produce the tables are available in the supplementary material (https://doi.org/10.6084/m9.figshare.9036884). A subscription to Scopus is required to replicate the research, except with updated citation counts, with the methods described above.

REFERENCES

Abramo, G., Cicero, T., & D'Angelo, C. A. (2011). Assessing the varying level of impact measurement accuracy as a function of the citation window length. *Journal of Informetrics*, *5*(4), 659–667.

Andersen, J.P., Schneider, J.W., Jagsi, R., & Nielsen, M.W. (2019). Gender variations in citation distributions in medicine are very small and due to self-citation and journal prestige. *eLife*, *8*, e45374. https://doi.org/10.7554/eLife.45374

Ashmos Plowman, D., & Smith, A. D. (2011). The gendering of organizational research methods: Evidence of gender patterns in qualitative research. *Qualitative Research in Organizations and Management: An International Journal*, *6*(1), 64–82.

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.

Carter, T. E., Smith, T. E., & Osteen, P. J. (2017). Gender comparisons of social work faculty using *h*-index scores. *Scientometrics*, *111*(3), 1547–1557.

Ceci, S. J., & Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences*, *108*(8), 3157–3162.

Didegah, F., & Thelwall, M. (2013). Which factors help authors produce the highest impact research? Collaboration, journal and document properties. *Journal of Informetrics*, *7*(4), 861–873.

Diekman, A. B., Steinberg, M., Brown, E. R., Belanger, A. L., & Clark, E. K. (2017). A goal congruity model of role entry, engagement, and exit: Understanding communal goal processes in STEM gender gaps. *Personality and Social Psychology Review*, *21*(2), 142–175.

European Commission. (2019). *She Figures 2018*. Luxembourg: Publications Office of the European Union.

Grant, J., Burden, S., & Breen, G. (1997). No evidence of sexism in peer review. *Nature*, *390*(6659), 438.

Grant, L., Ward, K. B., & Rong, X. L. (1987). Is there an association between gender and methods in sociological research? *American Sociological Review*, *52*(12), 856–862.

Guardian. (2005). Why women are poor at science, by Harvard president. https://www.theguardian.com/science/2005/jan/18/educationsgendergap.genderissues

Guglielmi, G. (2019). Eastern European universities score highly in university gender ranking. *Nature News*. https://doi.org/10.1038/d41586-019-01642-4

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, *75*(4), 800–802.

Jorstad, J., Starobin, S. S., Chen, Y., & Kollasch, A. (2017). STEM aspiration: The influence of social capital and chilly climate on female community college students. *Community College Journal of Research and Practice*, *41*(4–5), 253–266.

Kretschmer, H., & Kretschmer, T. (2013). Gender bias and explanation models for the phenomenon of women's discriminations in research careers. *Scientometrics*, *97*(1), 25–36.

Larivière, V., Desrochers, N., Macaluso, B., Mongeon, P., Paul-Hus, A., & Sugimoto, C. R. (2016). Contributorship and division of labor in knowledge production. *Social Studies of Science*, *46*(3), 417–435.

Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature News*, *504*(7479), 211.

Levitt, J., & Thelwall, M. (2013). Alphabetization and the skewing of first authorship towards last names early in the alphabet. *Journal of Informetrics*, *7*(3), 575–582.

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: a meta-analysis. *Psychological Bulletin*, *136*(6), 1123.

Mitchell, S. M., Lange, S., & Brus, H. (2013). Gendered citation patterns in international relations journals. *International Studies Perspectives*, *14*(4), 485–492.

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, *106*(1), 213–228.

Mongeon, P., Smith, E., Joyal, B., & Larivière, V. (2017). The rise of the middle author: Investigating collaboration and division of labor in biomedical research using partial alphabetical authorship. *PLOS ONE*, *12*(9), e0184601.

NCES. (2019a). Table 315.80. Full-time and part-time faculty and instructional staff in degree-granting postsecondary institutions, by race/ethnicity, sex, and program area: Fall 1998 and fall 2003, https://nces.ed.gov/programs/digest/d17/tables/dt17_315.80.asp

NCES. (2019b). 325: Trends in degrees by field. https://nces.ed.gov/programs/digest/d17/tables_3.asp#Ch3Sub25

Østby, G., Strand, H., Nordås, R., & Gleditsch, N. P. (2013). Gender gap or gender bias in peace research? Publication patterns and citation rates for *Journal of Peace Research*, 1983–2008. *International Studies Perspectives*, *14*(4), 493–506.

Othman, M., & Latih, R. (2006). Women in computer science: No shortage here! *Communications of the ACM*, *49*(3), 111–114.

Persson, O., Glänzel, W., & Danell, R. (2004). Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, *60*(3), 421–432.

Potthoff, M., & Zimmermann, F. (2017). Is there a gender-based fragmentation of communication science? An investigation of the reasons for the apparent gender homophily in citations. *Scientometrics*, *112*(2), 1047–1063.

Rousseeuw, P., Croux, C., Todorov, V., Ruckstuhl, A., Salibian-Barrera, M., Verbeke, T., & Maechler, M. (2019). *robustbase: Basic Robust Statistics*. R package version 0.92-8.

Stockard, J., Greene, J., Richmond, G., & Lewis, P. (2018). Is the gender climate in chemistry still chilly? Changes in the last decade and the long-term impact of COACh-sponsored workshops. *Journal of Chemical Education*, *95*(9), 1492–1499.

Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological Bulletin*, *135*(6), 859.

Tahamtan, I., Afshar, A. S., & Ahamdzadeh, K. (2016). Factors affecting number of citations: A comprehensive review of the literature. *Scientometrics*, *107*(3), 1195–1225.

Tellhed, U., Bäckström, M., & Björklund, F. (2017). Will I fit in and do well? The importance of social belongingness and self-efficacy for explaining gender differences in interest in STEM and HEED majors. *Sex Roles*, *77*(1–2), 86–96.

Thelwall, M. (2016). The discretised lognormal and hooked power law distributions for complete citation data: Best options for modelling and regression. *Journal of Informetrics*, *10*(2), 336–346.

Thelwall, M. (2017). Three practical field normalised alternative indicator formulae for research evaluation. *Journal of Informetrics*, *11*(1), 128–151.

Thelwall, M. (2018). Do females create higher impact research? Scopus citations and Mendeley readers for articles from five countries. *Journal of Informetrics*, *12*(4), 1031–1041.

Thelwall, M. (2020a). Gender differences in citation impact for 27 fields and six English speaking countries 1996–2014. *Quantitative Science Studies*, *1*(2), 599–617. https://doi.org/10.1162/qss_a_00038

Thelwall, M. (2020b). Large publishing consortia produce higher citation impact research but coauthor contributions are hard to evaluate. *Quantitative Science Studies*, *1*(1), 290–302.

Thelwall, M. (2020c). Female citation impact superiority 1996–2018 in six out of seven English-speaking nations. *Journal of the Association for Information Science and Technology*, *71*(8), 979–990.

Thelwall, M., Bailey, C., Makita, M., Sud, P. & Madalli, D. (2019a). Gender and research publishing in India: Uniformly high inequality? *Journal of Informetrics*, *13*(1), 118–131.

Thelwall, M., Bailey, C., Tobin, C. & Bradshaw, N. (2019b). Gender differences in research areas, methods and topics: Can people and thing orientations explain the results? *Journal of Informetrics*, *13*(1), 149–169.

Van Den Besselaar, P., & Sandström, U. (2016). Gender differences in research performance and its impact on careers: A longitudinal case study. *Scientometrics*, *106*(1), 143–162.

Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. (2011). Towards a new crown indicator: An empirical analysis. *Scientometrics*, *87*(3), 467–481.

Waltman, L. (2012). An empirical analysis of the use of alphabetical authorship in scientific publishing. *Journal of Informetrics*, *6*(4), 700–711.

Whittaker, R. J. (2008). Journal review and gender equality: A critical comment on Budden et al. *Trends in Ecology & Evolution*, *23*(9), 478–479.

Yohai, V. J. (1987). High breakdown-point and high efficiency robust estimates for regression. *Annals of Statistics*, *15*(2), 642–656.