The MIT Press

Do women undertake interdisciplinary research more than men, and do self-citations bias observed differences?

Henrique Pinheiro[iD], Matt Durning[iD], and David Campbell[iD]

Elsevier Canada, 4428 Boul. Saint-Laurent #500, Montreal, H2W 1Z5, Canada

ABSTRACT

Some studies have shown that women undertake interdisciplinary research more than men, whereas other studies have shown no difference by gender. Women have also been shown to self-cite less often than men, a difference at least partly mediated through differences in career stages and prior productivity. Existing evidence on gender-based differences in interdisciplinarity may therefore be biased. If interdisciplinarity is inferred from the disciplinary diversity of a paper's cited references, a greater share of self-citations by men could decrease their measured interdisciplinarity relative to women. Such biases could lead to erroneous conclusions, because after correcting for self-citations one might uncover that women participate in interdisciplinary research equally to, or less than, men. Given that funding for interdisciplinary research is gaining in importance, obtaining accurate measurements of interdisciplinarity by gender is highly relevant for funders so that they can take appropriate action(s) in leveling the playing field across gender. For instance, evidence suggests women are sometimes advised not to participate in interdisciplinary research due to the risk it represents for their career progression. This study shows that a paper's interdisciplinarity increases with the presence of female authors, accounting or not for self-citations in the interdisciplinarity measurement.

1. INTRODUCTION

Evidence suggests that junior female scientists are sometimes discouraged from pursuing interdisciplinary research (Smith-Doerr & Croissant, 2016). While interdisciplinary research does carry some risk, particularly to scientists in their early career stages, collaborative research efforts generally, as well as interdisciplinarity, may benefit the citation impact of the resulting publications (Beaudet, Campbell et al., 2014; Chen, Arsenault, & Larivière, 2015; Freeman & Huang, 2014). If women, relative to men, are more likely to perceive inter-disciplinary work as too high risk, and if funders do not take action to correct the potential consequences of such perceptions, the overall research ecosystem could be negatively impacted. For instance, women could end up with a lower propensity toward interdisciplinary research than men.

To assess the current situation, survey and bibliometric data can be used to compare the interdisciplinarity of male and female researchers. In both a survey by Rhoten and Pfirman (2007) and a large-scale bibliometric analysis in support of a recent evaluation of the Natural

Sciences and Engineering Research Council of Canada's Discovery Research Program (Science-Metrix, 2019), women were not found to be less interdisciplinary than men. Using publication data for roughly 10,000 NSERC awardees, the latter study showed that female researchers actually exhibited higher rates of highly interdisciplinary papers (those in the top 10%) than male researchers (11% vs. 9.2%).

This latter result, because it is contrary to the outcome expected from women being discouraged from taking part in interdisciplinary work, raised concerns that a measurement bias owing to women self-citing less than men (Andersen, Schneider et al., 2019; Chawla, 2016; King, Bergstrom et al., 2017) could have been at play. Indeed, interdisciplinarity was, as is often the case, inferred by the disciplinary diversity of the cited references of publications as in Leahey, Beckman, and Stanko (2017) and Porter and Rafols (2009). Accordingly, it is possible that a greater share of self-citations could decrease a paper's interdisciplinarity by concentrating references in the core disciplines of the authors, leading to a reduced balance of represented disciplines and the average distance between them. Testing if such a measurement bias is at play is important, as the bias could lead to erroneous conclusions regarding gender-based differences in the interdisciplinarity of research. If such a bias was present, one could find, after correcting for self-citations, that women do not differ from men in their propensity to perform interdisciplinary research or, perhaps, that they even participate less than men in such endeavors.

Mishra and colleagues (2018) recently uncovered that gender-based differences in self-citation rates in the medical sciences are, to a large extent, attributable to differences in productivity rather than gender. Women and men researchers with larger sets of prior publications have more material to self-cite, but women remain largely underrepresented in later career stages (assumed to be more productive) due to a greater attrition rate than men (Directorate-General for Research and Innovation, 2019; see Figure 6.1). This suggests that if self-citations do bias interdisciplinarity as measured from the disciplines represented among a paper's references, any higher interdisciplinarity score of women relative to men could be mediated through existing differences in the "average" productivity and/or career stages of women versus men. The above bias would thus also have implications when comparing early career to established researchers, regardless of their gender. In fact, in the recent evaluation of NSERC's Discovery Research Program mentioned above, early career researchers were also found to have slightly higher rates of highly interdisciplinary papers (top 10%) than established researchers (10.2% vs. 9.3%). This reinforces the possibility that self-citations introduce a downward bias in measuring interdisciplinarity. It is thus important to assess and resolve (if necessary) this risk prior to drawing any conclusions from potential interdisciplinary differences—for example, between genders. Such differences could indeed influence the strategies adopted by funding organizations in leveling the playing field for interdisciplinary research.

The goal of this study was thus to investigate the magnitude and direction of the relationship between the gender composition of a paper's coauthors and its interdisciplinarity, accounting for a potential measurement bias mediated by self-citations, among several other factors (e.g., number of authors, seniority, and prior publication output). This will help us assess whether, and if so by how much, researchers differ in their propensity to undertake interdisciplinary research as a function of their gender.

As an initial test of the impact of self-citations on a paper's interdisciplinarity, a paper's interdisciplinarity score computed (1) excluding its references directed toward one of its authors' prior publications (i.e., interdisciplinarity without self-citations) was compared to its score computed (2) using all its references (interdisciplinarity with self-citations). Subsequently,

Table 1. Total documents in data set

Data set	# of Documents
Scopus	29,037,077
Document type & year filter	14,179,694
ERA	8,760,081
Frascati fields of science	8,707,496

the interdisciplinary scores of papers, computed with and without self-citations, were compared across paper bins based on the share of female authors per paper—this to document the potential existence of an obvious gender bias mediated through self-citations. To address the study's core questions in a more robust manner, the relationship between gender and interdisciplinarity was then investigated using multivariate modeling, accounting for potential confounders (e.g., number of authors) and mediators (e.g., seniority, self-citations).

2. METHODS

2.1. Data Set

In this study, we made use of Scopus data for the 10-year period between 2010 and 2019. We considered only peer-reviewed publications (journal articles, reviews, and conference papers) classified across 174 mutually exclusive subfields as defined under Science-Metrix's journal-based classification of science[1]. This classification has recently been improved to reclassify, at the paper level, publications in multidisciplinary journals (e.g., *Nature, PNAS, Science, PLOS ONE*) (Rivest, Vignola-Gagné, & Archambault, 2021). This is the classification used throughout this paper in computing the interdisciplinarity scores of individual papers as well as in normalizing indicators.

As this work was supported by the European Commission, results were also reported by field of science according to the Frascati classification, which was used in *She Figures 2018* (Directorate-General for Research and Innovation, 2019). This was achieved by classifying Scopus papers by the Frascati fields of science, mainly by means of a direct conversion between the Science-Metrix subfields and the second-level Frascati fields. The only exception was for biotechnology in the former classification, which had multiple correspondences under the Frascati scheme (e.g., environmental biotech, industrial biotech). To resolve this issue, we reclassified biotechnology publications into the next best match per the Science-Metrix classification, thus obtaining a unique match to the Frascati scheme. This was achieved using the same AI algorithm developed to reclassify generalist journals within the Science-Metrix classification (Rivest et al., 2021). This algorithm takes into account the textual content of titles and abstracts, the citation links, and author information to predict the subfield most relevant to each paper. The analysis was further limited to papers where at least one author from a European Research Area (ERA) country was identified, for reasons detailed below. ERA countries consist of EU member states plus associated countries as defined in the *ERA Monitoring Handbook* (PPMI & Science-Metrix, 2018, p. 61). Table 1 provides the count of documents considered.

---

[1] https://science-metrix.com/?q=en/classification

Table 2. Number and share of authors by gender

Data set	Female authors		Men authors		Unassigned authors	
	#	%	#	%	#	%
Scopus	4,804,133	23.6	9,043,365	44.5	6,468,685	31.8
ERA	2,386,681	31.8	3,709,452	49.4	1,420,036	18.9

Over 8.7 million documents, with approximately 7.5 million distinct authors, were included in the analysis of gender and interdisciplinarity. Scopus includes an author identifier number (the AUID), which aims to aggregate each individual author's publications. This helps to identify self-citations, among other things, as the AUID identifies an author's full set of papers (removing false positives due to homonyms) even when there may be variations in the spelling or initialization of names. An assessment of the AUID showed that it produced reliable conclusions in a North American/European context when at least 1,000 authors were available per comparison group (Campbell & Struck, 2019). In a test set of 10,000 authors, their study also showed that the average recall and precision per author were, respectively, 98% and 96.9%; as the current study relies on millions of authors, we are confident that the use of AUIDs is offering robust results. However, similar information is lacking on its reliability in regions where homonyms are more prevalent (e.g., Asia). Additionally, inferring the gender of authors based on their names is less reliable in those same regions. We therefore limited this study to publications where at least one ERA member-country author was identified.

2.2. Gender

We genderized the author names in Scopus using the NamSor API[2]. As author names within an AUID may vary due to differences in spelling, typographical errors, or the presence of full name versus initials only (for which gender cannot usually be assigned), a method was devised to aggregate the predicted gender of the name variants of an AUID. For each name variant, the API generated a probability of the name being that of a man or woman. First, the average probability for each gender across name variants was computed for each AUID. Where this average for one gender exceeded 80%, the AUID was assigned to this gender. Second, if a name variant's probability exceeded 80% for one gender and the average of probabilities for that gender was between 65% and 80%, a gender was attributed. AUIDs that were assigned to multiple genders using this criterion (<0.05%), or for which a gender could not be attributed at all, were treated as unassigned. Across ERA countries for the period 2010–2019, approximately 18.9% of AUIDs were not assigned by the API using this gender assignation rule (GA #1). Further GA rules are presented in Section 2.5 to test the robustness of the findings obtained with different percentages of authors with an unassigned gender and different reliability levels of the assignation. Table 2 details the gender breakdown of AUIDs in Scopus and for the ERA data set.

Figure 1 presents the shares and trends in women's authorships relative to those of men in Scopus and across the Frascati fields of science for the ERA. In general, the rate of participation of women in scientific publications has been increasing across all fields of science for most countries. At the ERA level, the lowest share of female authorships is in Engineering and technology and the highest shares are in Social sciences and Humanities.
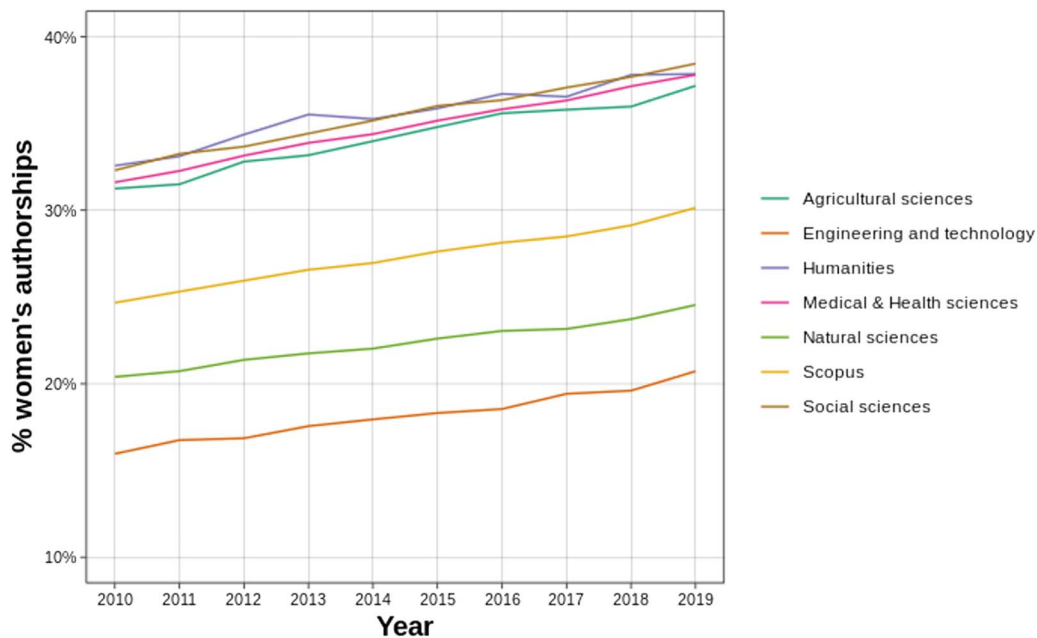
---

[2] https://namesorts.com/api/

Figure 1. Trend in share of women's authorships by Frascati field of science (2010–2019).

2.3. Self-Citations

To assess the rate (k) at which men self-cite relative to women, we used the men to women ratio of the mean self-citations per authorship as defined by King et al. (2017):

$$k = \left(\frac{s_m}{a_m}\right) / \left(\frac{s_w}{a_w}\right) \tag{1}$$

where s_m and s_w are the total self-citations made by men and women and a_m and a_w are the total authorships for men and women for a given country or region for a given period. An authorship is defined as an AUID–paper combination; if there are three authors on a paper, one man and two women, this amounts to one male authorship and two female authorships. An author-to-author self-citation is defined as an AUID citing one of its papers in a given paper (of which it has authorship); there can be multiple self-citations per authorship if the AUID self-cites more than one of its papers in that authorship.

As an example, take a paper (#1) with two authors, A and B, that cites two other papers, one (#2) by authors A and B, and another (#3) by author A and a third author C. The total number of author-to-author citations for paper #1 is equal to the sum, across its cited papers, of the product of its number of authors by that of the cited paper (i.e., $2 \times 2 + 2 \times 2 = 8$). In that example, there are three author-to-author self-citations (#1A–#2A, #1B–#2B, and #1A–#3A) out of eight author-to-author citations.

Because both the volume of citations and proportion of female authorships vary across fields of science, results were disaggregated by field.

2.4. Interdisciplinarity

Interdisciplinarity highlights instances where new knowledge (i.e., research publications) truly recombines *a priori* disparate knowledge (i.e., from diverse disciplines), assuming a paper's

references are a reliable indication that knowledge from the cited sources has been integrated in a novel way in the research project.

Following the work of Porter and Rafols (2009), interdisciplinarity was investigated using the Rao-Stirling index (RS) to quantify the diversity of integrated knowledge as represented in a publication's references. Each paper was assigned an interdisciplinarity score from 0 (i.e., completely following predominant citation patterns) to 1, the latter being extremely interdisciplinary (i.e., diverging completely from normal citation patterns in Scopus, integrating knowledge from areas that others do not), using the following formula:

$$\text{RS} = 1 - \sum_{i,j} s_{ij} p_i p_j = \sum_{i,j} d_{ij} p_i p_j \qquad (2)$$

where p_i and p_j are the respective proportions of references in subfields i and j in a paper's reference list (whereby $p_i p_j$ captures the variety and balance of represented subfields). The summation is taken over all cells of the subfield-by-subfield similarity matrix, accounting for all subfields in the Science-Metrix classification. s_{ij} is the cosine similarity between subfields i and j and captures how close (or distant, by taking $d_{ij} = 1 - s_{ij}$ as in the above formula for interdisciplinarity) the integrated subfields are in a given paper; the cosine similarity matrix between subfields is computed relying on the subfield cocitation network in a reference set of papers (here the whole of Scopus). The classification used to categorize publications by subfield can have an impact on the resulting interdisciplinarity scores. We refer the reader to the supplementary material of a prior publication by the authors for the rationale behind the selected classification as well as for more details on the computation of interdisciplinarity (Pinheiro, Vignola-Gagné, & Campbell, 2021).

In this paper, we have converted interdisciplinary scores into a binary variable identifying highly interdisciplinary publications (i.e., the top 10%). Because the score of a paper is in part dependent on how many references it includes, which varies across document types as well as across disciplines due to research practices and coverage issues in Scopus, and because interdisciplinary research has been shown to increase over time (Porter & Rafols, 2009), we identified the 10% most interdisciplinary papers by subfield, document type, and year. This procedure makes it impossible to study global differences in interdisciplinarity across scientific areas but enables identifying those publications that stand out relative to the "norm" in their respective subfield. As previously noted by Campbell, Deschamps et al. (2015), not using such a "normative" approach could otherwise lead to inappropriate comparisons across scientific areas due to coverage biases in bibliographic databases. The approach also enables comparisons across groups (e.g., between men and women) where interdisciplinarity is computed by aggregating the scores of publications over several years (e.g., 2010–2019). The groups being compared might not share the same yearly distribution of publications, which could otherwise advantage groups with a higher share of their publications in recent years.

In creating the binary variable identifying the 10% most interdisciplinary publications, fractioning of publications was used to ensure that exactly 10% of papers in Scopus fell in the top 10% by subfield, year, and document type. In Section 3.3, the binary variable thus obtained is used to compare interdisciplinarity across bins of the proportion of female authors on individual publications (with and without self-citations). In Section 3.4 on the multivariate regression models, the variable indicating whether a paper figures among the 10% most interdisciplinary is not fractioned and those papers tied on the edge of the top 10% and the 90% less interdisciplinary ones were classified among the top 10%.

Table 3. Total documents in data set after filters

Filter	# of documents
≥ 1 author gender identified	8,660,753
Interdisciplinarity score calculated	6,791,390

To account for differences over time and across subfields in a more transparent way, some of the regression models later introduced in Section 2.5 were rerun using the interdisciplinarity score as a continuous variable instead of a binary one identifying the top 10% most interdisciplinary papers. In these cases, a normalization was still present in the form of subfield and year fixed-effects (see supplementary material). The raw scores were also used in comparing the average interdisciplinarity of publications with and without self-citations in Section 3.2.

2.4.1. Interdisciplinarity with and without self-citations

Two series of paper-level and aggregated (e.g., ERA-level) interdisciplinarity scores were calculated: one considering all references made by papers in the database, and one where paper-to-paper self-citations were excluded. Paper-to-paper self-citations are those instances where at least one of the citing paper's AUIDs also appeared among the AUIDs of the cited papers. Note that self-citations were not removed in computing the subfields' proximity matrix to ensure comparability in the scores computed with and without self-citations at the paper level.

Once self-citations were removed, we kept only publications with at least five references with a known subfield in computing the raw and normalized scores with and without self-citations. The process of removing self-citations from the calculation resulted in fewer publications with enough references on which to base an interdisciplinarity score. We limited the comparison between scores based on self-citations and those without them to those papers for which the minimum threshold of references was met after the removal of self-citations.

In the end, the final data set used for this analysis was limited to only those papers for which a gender was assigned to at least one author and for which an interdisciplinarity score was computed. Table 3 details the total number of publications after the final filters were applied.

2.4.2. Interdisciplinarity by gender

As a first step in assessing the interdisciplinary nature of woman-led research, the share of female authors on a paper was calculated as the total number of AUIDs assigned to women over the total number of AUIDs where a gender was successfully identified. Papers were then separated into bins based on the proportional representation of female authors on a given paper, with bin 1 representing publications with between 0% and 10% female authors, and so on. The normalized share of the 10% most interdisciplinary papers was then computed, with and without author-to-author self-citations within each of these bins. In a second step, the magnitude and direction of the relationship between the number of female authors on a paper (while controlling for total number of authors) and its interdisciplinarity was further investigated using regression models enabling the integration of several control variables (Section 2.5).

Figure 2 shows the number of publications by bin within the ERA data set, by field of science and for all fields combined (i.e., in Scopus). In all fields, the lowest bin (0% to 10% female representation) dominated, ranging from approximately 22.6% in Medical & Health sciences to over 50% of all publications in Engineering and technology. Social sciences and
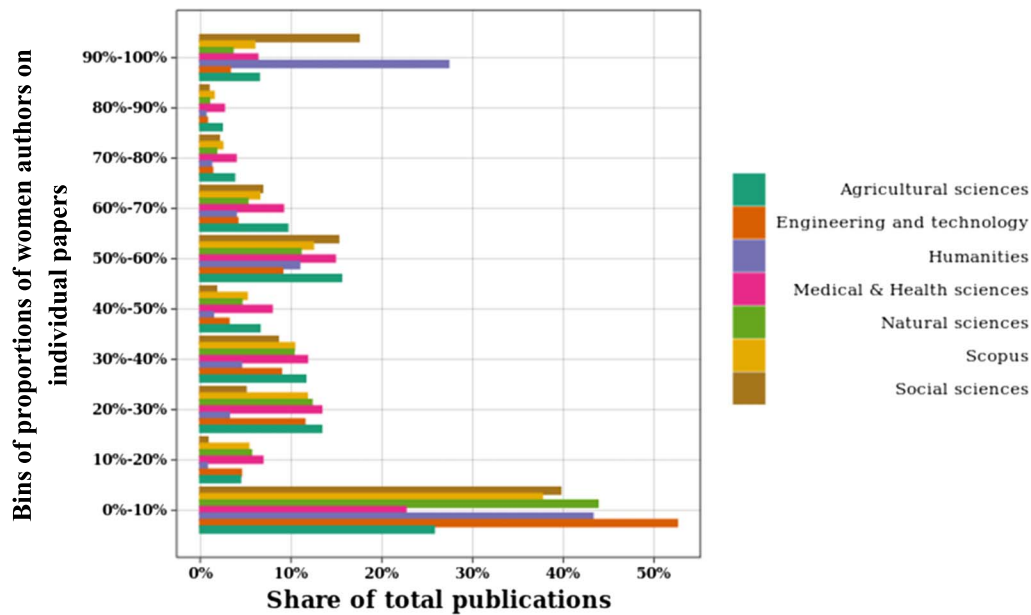
Figure 2. Share of publications per bin of female authors in Scopus overall and by Frascati field of science on ERA papers (2010–2019).

Humanities had by far the largest shares of publications in the top bin (90% to 100% female representation). Otherwise, it is worth noting that none of the fields of science had more than 50% of their publications with at least 50% of female authors. Of all fields of science, the most prominent ones from this perspective were Social sciences (42.6%), Humanities (44.1%), and the Medical & Health sciences (36.9%). This indicates that while the participation of women in research in these areas was among the highest (respectively 36%, 36%, and 35%), it was still less frequent for women than men to be all or most of the authors on a paper.

2.5. Multivariate Modeling

2.5.1. *Groups of papers binned by number of authors*

To further investigate the magnitude and direction of the relationship between the presence of female authors on a paper and its interdisciplinarity, multivariate analysis was performed using the R package "fixest" (Berge, 2018). Different models accounted for groups of papers defined per their number of authors (i.e., 1, 2, 3–5, 6–10, or 11–20 authors), which enabled a better specification of the gender variables included in the regression models. As an example, two dummy variables for the author's gender (one for female and one for unknown) were included in the model for single-authored papers. In this case, the coefficient for the variable "female author" estimated the interdisciplinary difference between women and men (the baseline). Similarly, the variable "unknown gender" estimated the interdisciplinary difference between unknown gender and men (the baseline). The same variables would not apply, for example, to papers having 11–20 authors. For those papers, the number of female (and unknown) authors was included among the predictors with the total number of authors included as a control variable. The approach of grouping papers by number of authors also enabled us to test for the presence and magnitude of the relationship between female authors and interdisciplinarity across papers with different numbers of authors. For example, gender differences could have been observed among single-author papers but not among those with

Table 4. Variables used in the various groups (by number of authors) of regression models

No. of authors	Variables for gender	Fixed-effects	Control variables
1	**Dummy variables:** Female author; Unknown gender.	Fixed-effects for subfield and year of publication	• Number of authors (for groups with 3–5 or more authors)
2	**Dummy variables:** All female; Mixed genders; Unknown genders (if the gender of any author was unknown).		**Only included in MS2:**
3–5	**Dummy variables:** One for each number of female authors in the paper (1–5 female authors).		• Number of references (complemented by squared and cubic terms) • Number of self-citations • Number of previous papers: count of papers previously published by any coauthor of a given paper. In the models with only two authors, the maximum and the minimum number of previous papers (across coauthors) was used, because it provided the number of papers of each coauthor. For the papers with more authors, having the minimum and the maximum number of papers would not capture all the previous output from the coauthors, so we used the total number of previous papers by the team. For single-author papers, these options would be equivalent.
6–10 and 10–20	Number of female authors; Number of unknown authors.		• Maximum seniority: highest number of years since first publication in Scopus among a paper's coauthors • Minimum seniority: smallest number of years since first publication in Scopus among a paper's coauthors

many authors. Table 4 (second column) summarizes the variables used for gender across the various groups of regressions.

2.5.2. Two model specifications per group of regressions

At first, the regression models were estimated using a simple model specification (MS1) including gender as a predictor of interdisciplinarity, the number of authors (for groups of regressions using papers with three or more authors) as a control variable, and subfield and year fixed-effects.

Including the number of authors as a control variable is important, as it may confound the effect of gender, expressed as the number of female authors (see Table 4), on interdisciplinarity. The number of authors influences the number of female authors, and larger teams are, to some degree, more likely than smaller ones to cross several disciplinary boundaries.

The subfield and year fixed-effects are also important as they mitigate the potential influence of subfields and years on the observed results. In the case of subfields, the fixed-effects control for differences in citation practices and database coverage across subfields that impact interdisciplinary measurements. At the same time, they eliminate an effect of women that would be mediated by gender differences in subfield preferences. For example, if women (or men) were relatively more active in subfields exhibiting a higher than average level of

interdisciplinarity, higher scores for women (or men) could be mediated by their higher presence in those subfields instead of reflecting a direct propensity for interdisciplinary research (the same goes for publication year, as both the presence of women in research and interdisciplinarity are increasing over time). This model specification was taken as our baseline for measuring the total effect of women on interdisciplinarity, disregarding differences owing to subfield/year of activity.

Subsequently, a second set of models (MS2) were estimated, adding further control variables to MS1. These were intended to control for other factors that could be partly mediating an effect of female authors on interdisciplinarity. Of particular relevance was a paper's number of author self-citations, which was used to assess whether part of the total effect of gender on interdisciplinarity may be mediated through a measurement bias. To properly assess the strength and direction of a potential bias introduced by self-citations, other controls included the number of prior publications by a paper's authors (measuring how prolific they are) and their seniority (as a team), as well as the paper's number of listed references (the causal chains linking these variables are further discussed in Section 2.5.3). Author characteristics were obtained via the Scopus AUIDs, which were, as previously mentioned, shown to produce reliable results (Campbell & Struck, 2019). See Table 4 for more details on the definition of these control variables for the various groups of regressions based on number of authors.
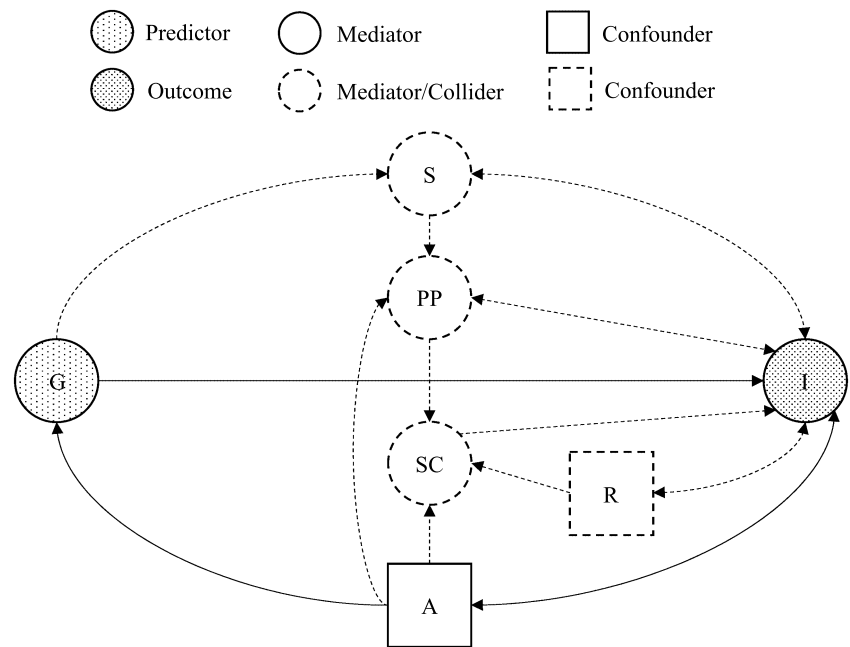
2.5.3. Interpretation of model specification 2

As introduced above, MS2 includes several control variables that might be confounding or mediating some of the "total" effects of women on interdisciplinarity. However, the complexity of plausible causal chains between the selected model variables, including the predictor, outcome, and control variables, is such that by including them all, we also run the risk of introducing a spurious correlation between women and interdisciplinarity due to a "collider" bias.

To help assess the risk of a collider bias in MS2, the most likely causal chains between the selected variables were depicted and analyzed using the R package "ggdag" (Malcolm, 2021). Figure 3 illustrates the complex relationship among the variables included in MS2, including the possibility of reverse causality (i.e., two-way causal relationship represented by bidirectional links) and collider bias.

The rationales underlying the depicted relationships are summarized below:

- **Gender (G) → Seniority (S) → Prior Publications (PP) → Self-Citations (SC) → Interdisciplinarity**: There is evidence in the literature showing that women publish less than men. However, this difference would, to a large extent, be attributable to differences in career lengths and dropout rates. After controlling for such factors, women and men would publish at a comparable annual rate (Huang, Gates et al., 2020). Accordingly, gender may indirectly influence the volume of prior publications via differences in seniority. Men, with a greater pool of prior publications relative to women (due to differences in career stages and career lengths), would thus have a greater pool of prior research to self-cite (Mishra et al., 2018). Assuming self-citations would belong to the main subfield of a publication's reported research, or to similar ones, this could induce a gender bias in interdisciplinarity (i.e., lower scores for men) as measured by the diversity of a paper's cited references. In this causal chain, seniority, the number of prior publications, and the number of self-citations may thus mediate a portion of the total effect of gender on interdisciplinarity that is due to a measurement bias.

Figure 3. Causal diagram depicting plausible relationships between the variables of model specification 2 (MS2). G: Gender (defined on the basis of the number of female authors); S: Seniority; PP: Prior publications; SC: Self-citations; R: References; A: Number of authors; I: Interdisciplinarity. Dashed nodes and links were removed in estimating the simplified version of MS2, that is MS1.

- **Seniority (S) ↔ Interdisciplinarity (I) and Prior Publications (PP) ↔ Interdisciplinarity (I)**: In addition to the above connections linking seniority and the number of prior publications to interdisciplinarity, there may be a direct relationship between each of these variables and interdisciplinarity. Prior research has shown that embarking on an interdisciplinary venture partly has to do with a researcher's self-motivation. In their efforts to uncover practical solutions to the complex problems facing modern societies, graduate students may be more open to pursuing unconventional paths than established researchers, as they are likely less entrenched in their disciplinary norms. That said, while talented early-career researchers may be attracted by the societal returns of an interdisciplinary career, they may be more frequently discouraged from pursuing such paths due to perceived risks for their career prospects (e.g., securing a tenure-track position) (Blackmore & Kandiko, 2011; Gewin, 2014; Rhoten & Parker, 2004). These perceptions may also be modulated by their prior achievements (e.g., it could be that the more publications a young researcher has, the lower the perceived risks). Seniority and/or the number of prior publications may thus also mediate part of a gender effect on interdisciplinarity by impacting a researcher's openness to this mode of research. On the other hand, the above causal relationships may be reversed, leading to two-way causal interactions. For instance, when recruiting team members, the interdisciplinary nature of a research question/project may lead a principal investigator to strike a balance between junior (more broad-based) and senior (more specialized and experienced) scientists. In such a case, both seniority and the number of prior publications would run the risk of inducing a collider bias in modeling the relationship between gender and interdisciplinarity.
- **Interdisciplinarity (I) ↔ References (R) → Self-Citations (SC)**: Another potential source of collider bias emerges from the inclusion of the number of self-citations and references in MS2. Recall that self-citations were included in the model to assess whether they

account for part of a gender effect on interdisciplinarity that would be attributable to a measurement bias, as well as to quantify the direction of this bias (hypothetically negative). In doing so, the number of references were also included as they positively relate to the number of self-citations and, per the above construction of the interdisciplinary indicators, would be expected to increase interdisciplinarity. In other words, the inclusion of the number of references is needed as it is confounding the direct link between self-citations and interdisciplinarity. However, because an interdisciplinary project may be expected to draw on a wider knowledge base than a monodisciplinary project, interdisciplinarity may also influence a paper's number of references (reverse causal link), converting the number of self-citations into a collider.

- **Interdisciplinarity (I) ↔ Authors (A) → Self-Citations (SC) (or → Prior Publications [PP])**: Much like the number of references, the inclusion of the number of authors may convert the number of self-citations into a collider. This is because an interdisciplinary question/project is likely to require input from a larger team, leading to an increase in the number of authors, which in turn may increase the number of self-citations (e.g., each author citing some of his or her prior work). Similarly, the inclusion of the number of authors may transform the number of prior publications into a collider (the more authors on a paper, the larger the authors' set of prior publications). As with the number of references, the number of authors may confound the relationship between self-citations and interdisciplinarity. Indeed, the larger a paper's number of authors, the more self-citations and the more disciplines it may contain.

- **Gender (G) ← Authors (A) ↔ Interdisciplinarity (I)**: Despite the downsides discussed above of including the number of authors as a control variable, including this variable is important, as the number of authors may otherwise confound the effect of gender on interdisciplinarity (as explained in Section 2.5.1). As interdisciplinary questions/projects might trigger larger teams, the causal relationship between gender and interdisciplinarity may be inverted, although we hypothesize that this driving force is likely less significant than the opposite scenario, whereby women would demonstrate greater preference/ability to work in an interdisciplinary context. For example, because female authors are, on average, younger, they may be more open to an interdisciplinary career and attracted by its potential to positively impact societies (Rhoten & Parker, 2004). Nevertheless, the direct relationship between gender and interdisciplinarity should be interpreted cautiously, emphasizing its strength over the direction of the causal link.

Testing robustness of results Findings from the multivariate modeling in the results section of this paper are those obtained using interdisciplinarity defined as a binary outcome variable called "highly interdisciplinary paper" (publications scored 1 if they figured among the top 10% most interdisciplinary in their subfield and year of publication, and 0 otherwise). To test the robustness of the study's results to changes in the mathematical formulation of interdisciplinarity, this variable has been computed using two diversity metrics: the Rao-Stirling Index (i.e., this paper's core interdisciplinarity metric; see Section 2.4) and the DIV* metric (Zhang & Leydesdorff, 2021). DIV* was originally developed to address some issues with the RS index and is formulated as:

$$\text{DIV}^* = n_c(1 - \text{Gini}_c) \sum_{i=1, j=1, i \neq j}^{i=n_c, j=n_c} d_{ij} / (n_c \times (n_c - 1)) \tag{3}$$

With DIV*, variety (n_c = number of cited subfields), balance ($1 - \text{Gini}_c$), and disparity ($\sum_{i=1, j=1, i \neq j}^{i=n_c, j=n_c} d_{ij} / (n_c \times (n_c - 1))$) are captured via independent components of diversity that are

subsequently combined, whereas with RS, variety and balance are captured as a single term (p_ip_j) *ex ante*, which is subsequently merged with disparity (d_{ij}) (see Section 2.4). Thus, DIV* may weigh variety, balance, and disparity more evenly than RS, with which more weight would be given to disparity. DIV* was thus an indicator of choice to test the reliability of this paper's results. Findings for both variants (RS and DIV*) are presented in the main body of the paper.

As noted in Section 2.2, approximately 18.9% of authors in this study's data set were not assigned a gender based on gender assignation rule #1 (GA #1). To further test the robustness of the findings to variation in the percentages of authors with an unassigned gender, and in the reliability of the assignation, results were produced using two additional GA rules (results for both rules are reported in the main body of the paper):

- **Relaxed rule (GA #2):** A gender was assigned to an author if the average probability for one gender across the name variants exceeded 50%. This led to a reduction in the share of authors with an unassigned gender from 18.9% with GA #1 to 9.0%. Under this rule, the unassigned names correspond to cases with no measured probability of corresponding to a woman or a man.
- **Stringent rule (GA #3):** A gender was assigned to an author if the average probability for one gender across the name variants exceeded 90% *or* if at least one of the name variants scored a probability for one gender exceeding 90%, with the average across the name variants for the corresponding gender being between 75% and 90%. This led to an increase in the share of authors with an unassigned gender from 18.9% with GA #1 to 27.4%.

Finally, to further test the robustness of the study's findings, the same models were tested using an alternative variable to express interdisciplinarity based on RS and DIV*. This variable consisted of the raw interdisciplinarity of papers (continuous variable instead of the above binary variable for highly interdisciplinary papers) computed with and without self-citations. Results based on these additional alternatives only appear in the supplementary materials and do not substantially change the results reported in the next section.

3. RESULTS

3.1. Self-Citations by Gender

Over the 2010–2019 period, the ratio at which men self-cited relative to women was 1.53, lower than the 1.71 figure of King et al. (2017) for 2000–2011. Still, it remained well above the expected value of 1 if rates were equal, reinforcing the notion that men do self-cite more frequently than women. Figure 4 depicts that ratio by field of science. Across all fields, men were more likely to self-cite than women, and the ratio at which they engaged in this behavior was similar for Social sciences, Agricultural sciences, and Medical & Health sciences. In Humanities, the field closest to gender parity in terms of total number of authors by gender, the gap was considerably smaller. In Natural sciences, men appear to have had a significantly higher tendency to self-cite than women.

These results suggest that the interdisciplinarity of a paper measured as the disciplinary diversity of its cited references may be biased downward for a single-author paper by a man compared to a woman. If self-citations were more likely to belong to the paper's core subfields, a greater share of self-citations could lead to a smaller disciplinary diversity of the cited references by reducing the balance of represented disciplines and the average distance
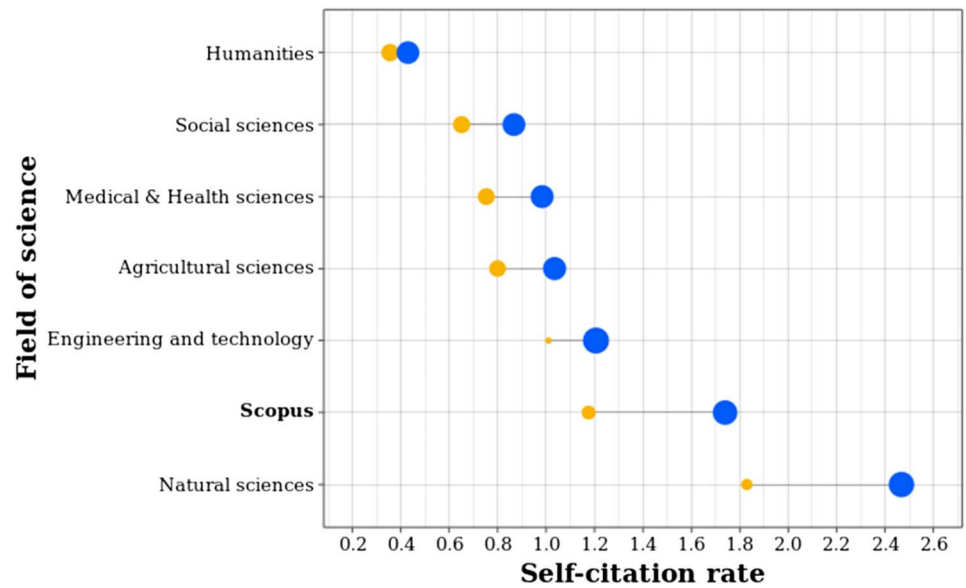
Figure 4. Ratio of male to female self-citation rates by field of science for ERA papers (2010–2019). The size of the bubble (yellow for women, blue for men) is proportional to the share of authors in a given field for each gender.

between them. In the context of cross-disciplinary collaboration, a paper's set of self-citations may represent a mixed bag of disciplines corresponding to the background of the contributing authors, which may confound a self-citation bias in measuring interdisciplinarity. This is one reason for the number of coauthors being included as a confounder in modeling the relationship between self-citations and interdisciplinarity (see Section 3.4).

3.2. Self-Citations and Interdisciplinarity

To test the effect of self-citations on the interdisciplinarity of publications, their raw (nonnormalized) interdisciplinarity was computed after removing their self-citations. The comparison between their scores with and without self-citations provided an initial test for the impact of self-citations on a paper's interdisciplinarity.

At the paper level, the average difference between both measurements of interdisciplinarity (with minus without self-citations) was neutral, or at most very slightly positive, whereas we expected it to be negative (average difference = 0.0021). When the subfield and year of publications were used as the unit of analysis (i.e., taking the average difference of the average interdisciplinarity across subfields and years), the average difference remained almost unchanged (0.0026).

However, using this approach, we are effectively comparing interdisciplinarity using two different sets of references, one (with self-citations) being larger than the other (without self-citations). This may lead to a comparability issue because the length of a reference list can impact its interdisciplinary scores in several ways. For example, the longer a reference list, the higher the odds of having a greater variety of represented subfields. This may outweigh a reduction in the balance and disparity of represented subfields from the self-citations. This may hold true even if the additional subfields from the self-citations are closely related to those of the nonself-citations. Diversity metrics are complex indicators capturing the variety of represented disciplines as well as the distance and balance between them. This makes it difficult to apprehend the impact of a paper's features on the resulting score.

To be fair, the approach would have to control for the difference in a publication's number of references, with and without self-citations. While this cannot be achieved using the simple difference approach presented here, the regression models presented in Section 3.4 were designed to account for this difference by incorporating a paper's total number of references and its number of self-citations as controls. With this approach, the effect of self-citations on interdisciplinarity (their regression coefficient) can now be interpreted as though a paper's total number of references was held constant.

3.3. Interdisciplinarity and Gender

Figure 5 shows the share of ERA papers among the 10% most interdisciplinary papers in Scopus (2010–2019), with and without self-citations, across bins of the share of female authors per paper. It demonstrates that with or without self-citations, more women being represented on a paper tends to correspond with higher levels of interdisciplinarity. This finding was consistent for bins of publications with varying number of authors (data not shown). The role of the number of authors as a potential confounder of the relationship between gender and interdisciplinarity is further investigated in Section 3.4.

In Figure 5, the scores with self-citations are generally close to or slightly higher than those without self-citations. This is most likely because the group of ERA papers experienced a relative increase in interdisciplinarity compared to non-ERA papers, with self-citations included. A case in point is that over the 2010–2019 period, the average disciplinary diversity of authors—a measure that captures the diversity in the disciplinary background of coauthors—was 5% higher for copublications with ERA authors than for copublications without ERA authors (1.03 vs. 0.98). Accordingly, the pool of potential self-citations that was available to the coauthors of ERA copublications was, on average, very likely to be more diversified than the pool available to the coauthors of non-ERA copublications. In
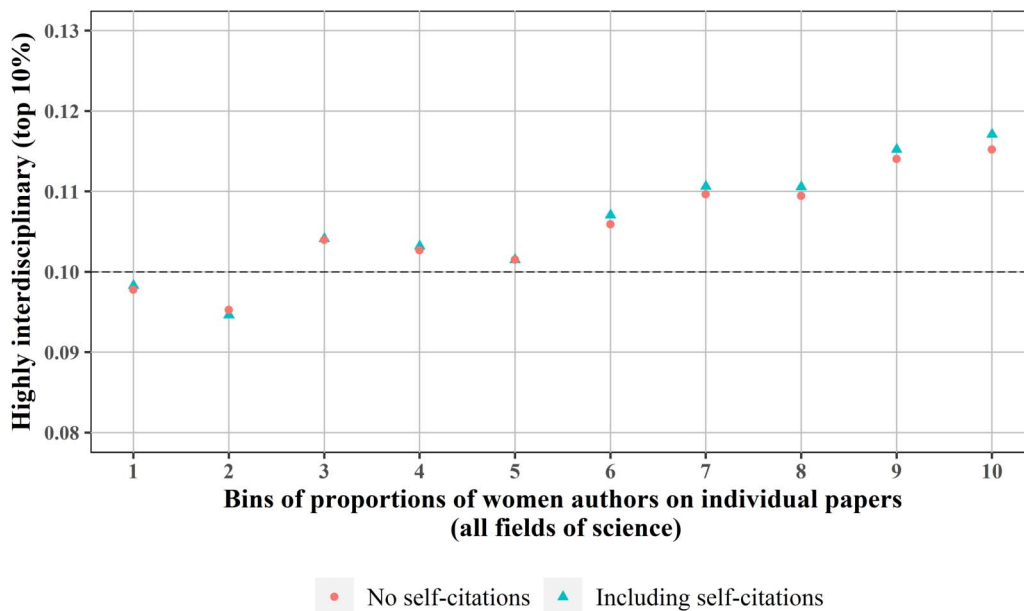
Figure 5. Highly interdisciplinary (top 10%) ERA papers by share of female authors across all fields of science (2010–2019).

turn, with self-citations included, ERA papers would have appeared more frequently, relative to non-ERA papers, among the 10% most interdisciplinary publications in the world (i.e., ERA plus non-ERA) than without self-citations included. Readers are referred to Pinheiro et al. (2021) for the method underlying the computation of the average disciplinary diversity of authors.

3.4. Multivariate Modeling of the Relationship Between Gender and Interdisciplinarity

Figure 5 does not account for potential confounder(s)/mediator(s) that could have driven the overall increase in interdisciplinarity for higher shares of women as coauthors. As explained in Section 2.5.3, the size of the research team could confound an effect of gender (as defined in Table 4) on interdisciplinarity, although several other factors could mediate such an effect (Figure 3). Among them, the number of self-citations would mediate a measurement bias owing to women self-citing less than men. A model specification (MS2) including all the confounders and mediators identified in Figure 3 was thus used to test whether the apparent relationship between gender and interdisciplinarity exhibited in Figure 5 is still in place after controlling for such variables.

The inclusion of such controls in MS2 unfortunately led to the inclusion of potential colliders (Figure 3) that could bias the measured relationship between gender and interdisciplinarity. In this context, the simpler model specification (MS1), which only controlled for a key confounder (i.e., a paper's number of authors), served the purpose of assessing whether the relationship between gender and interdisciplinarity remained significant upon exclusion of the potential colliders (i.e., seniority, prior publications, and self-citations) and the number of references (i.e., a confounder of the effect of self-citations on interdisciplinarity).

Table 5 summarizes the results for the effect on interdisciplinarity of gender, under both model specifications, and self-citations, under MS2. Recall that both model specifications included subfield and year fixed-effects and that the results summarized in Table 5 are based on GA #1 (the intermediary option between the stringent and relaxed rules), the RS index to measure interdisciplinarity, as well as interdisciplinarity defined as a binary outcome (1 = paper among the 10% most highly interdisciplinary papers; 0 = otherwise).

Tables 6–10, and the supplementary material, provide detailed results of the logistic regressions summarized in Table 5 plus the results of the robustness tests described in Section 2.5. For the two groups of papers with more than 6 authors (6–10 and 11–20 authors), all models included an interaction term between the number of female authors and the number of authors because the effect of the former variable may be higher in smaller teams. Additionally, they included quadratic terms for the number of female authors alone and interacted with the total number of authors. This is to account for possible nonlinear effects whereby the first woman added to a research team might have a larger impact on interdisciplinarity than the 10th woman added to a team.

The coefficients for these three terms are omitted in Table 9 and Table 10 (they are available in the supplementary material) as their effect was integrated into the reported coefficient for the number of female authors and the number of authors. Also note that in these tables, the effect of each additional female author is conditional on the number of women already figuring in a research team and on the size of the team. Accordingly, the reported odds ratios in Table 9 (6–10 authors) are for seven-author papers (the average in this group being 7.3) with two women already figuring in the team (the average being 2.13). In Table 10 (11–20 authors), they are for 13-author papers (the average being 13.4) with four women already in the team (the average being 4.11).

Table 5. Summary of logistic regressions on the link between gender (GA #1) and interdisciplinarity (RS) accounting or not for self-citations

Group of papers	Interpretation of the gender variables' coefficients*	Interpretation of the self-citation variable's coefficient*
1: Single-authored publications	Publications from women are, respectively in MS1 and MS2, 11.1% and 5.7% more likely than those of men to be among the 10% most interdisciplinary publications.	After controlling for a paper's number of references and number of coauthors, two confounders of the effect of self-citations on interdisciplinarity, self-citations exert a slightly negative and statistically significant effect on interdisciplinarity in nearly all groups of regressions. Each additional self-citation is associated with a decrease in the range from −0.5% to −1.5% for groups of papers with 1 to 10 coauthors (i.e., groups 1–4 in the first column of this table). Only for papers with 11–20 coauthors is there no measurable effect.
2: Publications with two authors	Publications authored by two women are, respectively in MS1 and MS2, 20.9% and 12.7% more likely than those of two men to belong to the top 10% most interdisciplinary publications; publications combining one woman and one man are, respectively in MS1 and MS2, 12.7% and 9.2% more likely to figure among the top 10% (compared with "all male" publications).	
3: Publications having 3–5 authors	"All women" publications are, respectively in MS1 and MS2, 24.8% and 11.7% more likely to be among the top 10 most interdisciplinary ones (compared with all-male papers). The effect ranges from 10.6% to 24.8% in MS1, and from 8.9% to 13.9% in MS2, considering all possible numbers of women in the team (from one to five).	
4: Publications having 6–10 authors	For a seven-author paper with two women already figuring in the team, an additional woman author is associated with, respectively in MS1 and MS2, an increase of 6.9% and 5.5% in the probability of the paper being among the top 10%. For seven-author papers under MS2, the effect varies from 9.5% for the first woman added to the research team to 3.3% for the fourth woman added. After this point (i.e., research teams with at least four women) no effect on interdisciplinary was associated with new female authors being added to the team.	
5: Publications having 11–20 authors	For a 13-author paper with three women already figuring in the team, an additional woman author is associated with an increase of 4.9% in the probability of the paper being among the top 10% using MS1. There was no robust association using MS2.	

* The interpretation is made relative to a baseline probability of 10% of highly interdisciplinary papers. The effects (for both gender and self-citations) reported in this table are even more pronounced when using DIV* as the diversity metric for measuring interdisciplinarity. For instance, the effect of self-citations with 11–20 coauthors then becomes negative and significant.

To assess the practical relevance of observed differences in interdisciplinarity between women and men, the estimated odds ratios (Tables 6 to 10) were "translated" into "change in probability of a paper belonging to the group of 10% most interdisciplinary papers" in Table 5. As an example, in the case of single-authored papers under MS1, a publication would be 11.1% more likely to figure among the most interdisciplinary if authored by a woman, using 10% as a reference (i.e., the expected probability of a paper belonging to the group of 10% most interdisciplinary papers). This means that the probability of a publication authored by one woman would be 11.11% (1.111 × 10%), relative to a 10% baseline for a paper authored by one man. The same approach applies to the interpretation of the effect of self-citations.

Table 6. Logistic regression models for highly interdisciplinary papers (single-author papers). Dependent variable: highly interdisciplinary publication (among 10% most interdisciplinary in the subfield and year of publication)

	Gender assignation (GA) #1			GA #2	GA #3
	RS$_{MS1}$	RS	DIV*	RS	RS
Female author	1.125***	1.064*	1.115***	1.062*	1.060*
	[1.062; 1.192]	[1.005; 1.126]	[1.053; 1.180]	[1.006; 1.122]	[1.000; 1.123]
Unknown gender	0.983	0.928*	1.001	0.797**	0.982
	[0.923; 1.046]	[0.868; 0.993]	[0.942; 1.063]	[0.677; 0.939]	[0.935; 1.032]
N references		1.001*	1.012***	1.001*	1.001*
		[1.000; 1.002]	[1.010; 1.014]	[1.000; 1.002]	[1.000; 1.002]
N self-citations		0.983***	0.978***	0.983***	0.983***
		[0.975; 0.991]	[0.970; 0.986]	[0.975; 0.991]	[0.975; 0.991]
N previous papers		0.998***	0.999*	0.998***	0.998***
		[0.997; 0.999]	[0.998; 1.000]	[0.997; 0.999]	[0.997; 0.999]
Seniority		0.994*	0.994'	0.994*	0.994*
		[0.988; 0.999]	[0.989; 1.000]	[0.988; 0.999]	[0.989; 0.999]
Fixed-effects:					
year	Yes	Yes	Yes	Yes	Yes
subfield	Yes	Yes	Yes	Yes	Yes
S.E. type	year/subfield	year/subfield	year/subfield	year/subfield	year/subfield
Observations	562,876	559,361	559,361	559,361	559,361
Squared Cor.	0.012	0.015	0.016	0.015	0.015
Pseudo R2	0.017	0.020	0.026	0.020	0.020
BIC	373,346	369,418	331,566	369,402	369,441

The coefficients reported in this table refer to the odds ratios. 95% confidence intervals are reported in brackets. The coefficients and clustered standard errors are reported in the supplementary material. Significance levels: $'p < 0.1$; $*p < 0.05$; $**p < 0.01$; $***p < 0.001$.

The presence of women on papers was associated with a higher probability of papers figuring among the most interdisciplinary in their subfield and year for the groups of papers involving up to 10 authors (Table 5). For these four groups of papers, the measured effects of gender on interdisciplinarity were robust to changes in the model specification, the gender assignation rule, the diversity metrics for measuring interdisciplinarity, and the variable type used to express interdisciplinarity (i.e., as a binary or continuous outcome variable) (Tables 6–9; also see Tables S1–S4 and S6–S9 in the supplementary material). The regression coefficients for the defined gender variables were nearly always pointing to a positive and statistically significant effect of women on interdisciplinarity. This is not necessarily surprising given the high number of observations underlying the estimation of parameters in these models. The only exception related to the coefficients (or odds ratios) not being statistically

Table 7. Logistic regression models for highly interdisciplinary papers (publications having only two authors). Dependent variable: highly interdisciplinary publication (among 10% most interdisciplinary in the subfield and year of publication)

	Gender assignation (GA) #1			GA #2	GA #3
	RS$_{MS1}$	RS	DIV*	RS	RS
All female	1.238***	1.144**	1.296***	1.141**	1.143**
	[1.127; 1.359]	[1.048; 1.248]	[1.195; 1.405]	[1.048; 1.242]	[1.049; 1.247]
Mixed genders	1.134***	1.103***	1.171***	1.010***	1.102***
	[1.074; 1.197]	[1.048; 1.162]	[1.113; 1.232]	[1.047; 1.155]	[1.044; 1.164]
Unknown genders	1.038'	1.000	1.052**	0.939	1.023
	[0.997; 1.080]	[0.962; 1.038]	[1.016; 1.089]	[0.867; 1.017]	[0.989; 1.058]
N references		0.999	1.018***	0.999	0.999
		[0.997; 1.001]	[1.009; 1.012]	[0.997; 1.001]	[0.998; 1.001]
N self-citations		0.986**	0.976***	0.986**	0.986**
		[0.978; 0.995]	[0.969; 0.983]	[0.978; 0.995]	[0.978; 0.995]
Max. previous papers		0.999***	1.000	0.999***	0.999***
		[0.999; 1.000]	[0.999; 1.000]	[0.999; 1.000]	[0.999; 1.000]
Min. previous papers		0.999*	1.000	0.999*	0.999*
		[0.998; 1.000]	[0.999; 1.000]	[0.998; 1.000]	[0.998; 1.000]
Max. seniority		0.984***	0.986***	0.984***	0.984***
		[0.980; 0.987]	[0.982; 0.990]	[0.980; 0.987]	[0.980; 0.987]
Min. seniority		0.999	0.995**	0.999	0.999
		[0.995; 1.003]	[0.992; 0.999]	[0.995; 1.003]	[0.995; 1.003]
Fixed-effects:					
year	Yes	Yes	Yes	Yes	Yes
subfield	Yes	Yes	Yes	Yes	Yes
S.E. type	year/subfield	year/subfield	year/subfield	year/subfield	year/subfield
Observations	1,067,909	1,060,901	1,060,901	1,060,901	1,060,901
Squared Cor.	0.005	0.008	0.011	0.008	0.008
Pseudo R2	0.008	0.011	0.019	0.011	0.011
BIC	695,421	688,114	637,425	688,092	688,159

The coefficients reported in this table refer to the odds ratios. 95% confidence intervals are reported in brackets. The coefficients and clustered standard errors are reported in the supplementary material. Significance levels: $'p < 0.1$; $*p < 0.05$; $**p < 0.01$; $***p < 0.001$.

Table 8. Logistic regression models for highly interdisciplinary papers (publications having 3 to 5 authors). Dependent variable: highly interdisciplinary publication (among 10% most interdisciplinary in the subfield and year of publication)

	Gender assignation (GA) #1			GA #2	GA #3
	RS$_{MS1}$	RS	DIV*	RS	RS
4 authors	1.019	1.071***	1.086***	1.069***	1.072***
	[0.9856; 1.0534]	[1.0391; 1.1034]	[1.0587; 1.1140]	[1.0377; 1.1011]	[1.0403; 1.1047]
5 authors	1.01	1.114***	1.149***	1.110***	1.116***
	[0.9516; 1.0728]	[1.0541; 1.1778]	[1.0957; 1.2053]	[1.0512; 1.1726]	[1.0552; 1.1797]
1 female author	1.119***	1.100***	1.184***	1.082***	1.098***
	[1.0642; 1.1775]	[1.0482; 1.1544]	[1.1370; 1.2329]	[1.0388; 1.1264]	[1.0445; 1.1531]
2 female authors	1.206***	1.157***	1.330***	1.142***	1.150**
	[1.1032; 1.3173]	[1.0641; 1.2578]	[1.2323; 1.4349]	[1.0594; 1.2314]	[1.0566; 1.2513]
3 female authors	1.238***	1.155**	1.429***	1.140*	1.154*
	[1.1042; 1.3886]	[1.0375; 1.2858]	[1.2954; 1.5757]	[1.0301; 1.2621]	[1.0335; 1.2884]
4 female authors	1.244**	1.131′	1.435***	1.113	1.139′
	[1.0707; 1.4457]	[0.9834; 1.3012]	[1.2670; 1.6257]	[0.9765; 1.2697]	[0.9903; 1.3092]
5 female authors	1.284**	1.132	1.474***	1.108	1.154′
	[1.0890; 1.5128]	[0.9692; 1.3216]	[1.2973; 1.6753]	[0.9534; 1.2888]	[0.9759; 1.3654]
1 unknown gender	0.977′	0.967**	0.987	0.96	0.979*
	[0.9543; 1.0009]	[0.9448; 0.9895]	[0.9643; 1.0098]	[0.9136; 1.0094]	[0.9606; 0.9982]
2 unknown genders	0.958′	0.917***	0.955*	0.917*	0.954**
	[0.916; 1.001]	[0.881; 0.955]	[0.917; 0.994]	[0.853; 0.986]	[0.923; 0.986]
3 unknown genders	0.943′	0.859***	0.940*	0.728***	0.906***
	[0.882; 1.007]	[0.812; 0.908]	[0.890; 0.994]	[0.653; 0.811]	[0.871; 0.942]
4 unknown genders	1.044	0.892′	0.971	0.723**	0.896**
	[0.926; 1.177]	[0.794; 1.002]	[0.870; 1.083]	[0.578; 0.905]	[0.833; 0.964]
5 unknown genders	0.873	0.677***	0.661***	0.819	0.889′
	[0.683; 1.115]	[0.543; 0.845]	[0.549; 0.795]	[0.539; 1.245]	[0.789; 1.001]
N references		1.000	1.011***	0.999	0.999
		[0.998; 1.001]	[1.010; 1.013]	[0.998; 1.001]	[0.998; 1.001]
N self-citations		0.982***	0.976***	0.983***	0.982***
		[0.977; 0.988]	[0.971; 0.981]	[0.977; 0.988]	[0.977; 0.988]
N previous papers		0.999***	0.9998*	0.999***	0.999***
		[0.999; 1.000]	[0.9995; 1.0000]	[0.999; 1.000]	[0.999; 1.000]

Table 8. *(continued)*

	Gender assignation (GA) #1			GA #2	GA #3
	RS$_{MS1}$	RS	DIV*	RS	RS
Max. seniority		0.981***	0.986***	0.981***	0.981***
		[0.976; 0.986]	[0.981; 0.991]	[0.977; 0.986]	[0.976; 0.986]
Min. seniority		0.996**	0.991***	0.996*	0.996**
		[0.992; 0.999]	[0.987; 0.994]	[0.993; 0.999]	[0.992; 0.999]
Fixed-effects:					
year	Yes	Yes	Yes	Yes	Yes
subfield	Yes	Yes	Yes	Yes	Yes
S.E. type	year/subfield	year/subfield	year/subfield	year/subfield	year/subfield
Observations	3,281,609	3,259,248	3,259,248	3,259,248	3,259,248
Squared Cor.	0.002	0.005	0.011	0.005	0.005
Pseudo R2	0.003	0.007	0.017	0.007	0.007
BIC	2,222,909	2,198,571	2,073,138	2,198,963	2,198,662

The coefficients reported in this table refers to the odds ratios. 95% confidence intervals are reported in brackets. The coefficients and clustered standard errors are reported in the supplementary material. Significance levels: '$p < 0.1$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

significant, although still positive, for two out of five gender variables in the third group of regressions for publications with 3–5 coauthors, and only in the following cases (Table 8):

- the variable five female authors using RS and MS2 with both the moderate and relaxed (less precise) gender assignation rule (i.e., GA #1 and #2), and
- the variable four female authors using RS and MS2 with the relaxed (less precise) gender assignation rule (GA #2).

Regarding the models involving papers with 11–20 authors, the relationship between gender and interdisciplinarity was not statistically significant for MS2 when the indicator variable for papers among the 10% most interdisciplinary publications was based on the RS index (Table 5). While a positive association between gender and interdisciplinarity was still observed in this group of papers using DIV* (Table 10), or when the raw scores for the RS index were used (Table S10 in the supplementary material), the association between the number of female authors and interdisciplinarity is less robust than observed in groups of papers with smaller research teams.

Recall that for the two groups of papers with more than six authors (6–10 and 11–20 authors), the reported odds ratios for the number of female authors in Tables 9 and 10 are conditional on the number of women already figuring in a research team and on the size of the team. This is due to the inclusion of quadratic terms for the number of female authors alone and interacted with the total number of authors. These quadratic terms were introduced to account for a potential nonlinear relationship whereby the effect of an additional woman on interdisciplinarity would depend on the number of women already included in the research

Table 9. Logistic regression models for highly interdisciplinary papers (publications having 6 to 10 authors†). Dependent variable: highly interdisciplinary publication (among 10% most interdisciplinary in the subfield and year of publication)

	Gender assignation (GA) #1			GA #2	GA #3
	RS$_{MS1}$	RS	DIV*	RS	RS
N authors (total)	0.957***	0.994	1.0124	0.9941	0.994
	[0.938; 0.977]	[0.973; 1.016]	[0.994; 1.031]	[0.973; 1.015]	[0.972; 1.017]
N female authors	1.077***	1.061**	1.139***	1.0551**	1.056**
	[1.032; 1.123]	[1.019; 1.104]	[1.099; 1.180]	[1.016; 1.095]	[1.015; 1.097]
N unknown gender	0.980*	0.962***	0.991	0.9852	0.974***
	[0.963; 0.998]	[0.944; 0.979]	[0.971; 1.010]	[0.958; 1.014]	[0.959; 0.988]
N references		1.000	1.013***	1.000	1.000
		[0.999; 1.002]	[1.011; 1.015]	[0.998; 1.001]	[0.999; 1.002]
N self-citations		0.994*	0.983***	0.9944*	0.994*
		[0.988; 1.000]	[0.978; 0.989]	[0.989; 1.000]	[0.988; 0.999]
N previous papers		0.999***	0.9997**	0.9995***	0.9995***
		[0.999; 1.000]	[1.000; 1.000]	[0.9993; 0.9997]	[0.9993; 0.9997]
Max. seniority		0.979***	0.987***	0.9815***	0.979***
		[0.971; 0.987]	[0.980; 0.994]	[0.974; 0.989]	[0.971; 0.987]
Min. seniority		0.991***	0.988***	0.9925**	0.991***
		[0.986; 0.995]	[0.983; 0.993]	[0.988; 0.997]	[0.986; 0.995]
Fixed-effects:					
year	Yes	Yes	Yes	Yes	Yes
subfield	Yes	Yes	Yes	Yes	Yes
S.E. type	year/subfield	year/subfield	year/subfield	year/subfield	year/subfield
Observations	1,850,767	1,840,752	1,840,752	1,840,752	1,840,752
Squared Cor.	0.006	0.009	0.018	0.008	0.009
Pseudo R2	0.008	0.012	0.026	0.012	0.012
BIC	1,214,649	1,204,412	1,226,743	1,205,039	1,204,500

The coefficients reported in this table refers to the odds ratios. 95% confidence intervals are reported in brackets. The coefficients and clustered standard errors are reported in the supplementary material.

† The reported coefficient for "N female authors" refers to the effect of moving from two to three female authors in a team of seven authors (see Figure 6 for the effect with different starting number of female authors). It integrates the effect of omitted terms: an interaction term between the number of female authors and the total number of authors as well as quadratic terms for the number of female authors alone and interacted with the total number of authors (see Table S11 in the supplementary material for more details; refer to the above description of results for more details).

Significance levels: ′$p < 0.1$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Table 10. Logistic regression models for highly interdisciplinary papers (publications having 11 to 20 authors†). Dependent variable: highly interdisciplinary publication (among 10% most interdisciplinary in the subfield and year of publication)

	Gender assignation (GA) #1			GA #2	GA #3
	RS$_{MS1}$	RS	DIV*	RS	RS
N authors (total)	0.951***	0.982*	0.969***	0.981*	0.983'
	[0.936; 0.967]	[0.965; 0.998]	[0.955; 0.983]	[0.963; 0.998]	[0.967; 1.000]
N female authors	1.055*	1.035'	1.085***	1.035'	1.03
	[1.010; 1.102]	[0.994; 1.078]	[1.044; 1.128]	[0.994; 1.078]	[0.992; 1.068]
N unknown gender	0.993	0.972*	1.007	0.981	0.978*
	[0.965; 1.021]	[0.945; 0.999]	[0.977; 1.036]	[0.915; 1.051]	[0.959; 0.998]
N references		1.001	1.013***	1.001	1.002'
		[1.000; 1.003]	[1.011; 1.014]	[1.000; 1.003]	[1.000; 1.003]
N self-citations		0.999	0.986***	0.999	0.999
		[0.993; 1.005]	[0.981; 0.992]	[0.993; 1.005]	[0.993; 1.005]
N previous papers		0.9996***	0.9998***	0.9996***	0.9996***
		[0.9995; 0.9997]	[0.9997; 0.9999]	[0.9995; 0.9997]	[0.9995; 0.9997]
Max. seniority		0.961***	0.978***	0.969***	0.961***
		[0.947; 0.975]	[0.967; 0.988]	[0.955; 0.983]	[0.947; 0.975]
Min. seniority		0.978***	0.962***	0.979***	0.976***
		[0.968; 0.985]	[0.953; 0.971]	[0.970; 0.989]	[0.968; 0.985]
Fixed-effects:					
Year	Yes	Yes	Yes	Yes	Yes
subfield	Yes	Yes	Yes	Yes	Yes
S.E. type	year/subfield	year/subfield	year/subfield	year/subfield	year/subfield
Observations	408,544	407,319	407,374	407,319	407,319
Squared Cor.	0.017	0.020	0.031	0.019	0.020
Pseudo R2	0.026	0.031	0.044	0.030	0.031
BIC	214,347	212,669	248,554	212,802	212,670

The coefficients reported in this table refers to the odds ratios. 95% confidence intervals are reported in brackets; The coefficients and clustered standard errors are reported in the supplementary material.

† The reported coefficient for "N female authors" refers to the effect of moving from four to five female authors in a team of 13 authors (see Figure 6 for the effect with different starting number of female authors). It integrates the effect of omitted terms: an interaction term between the number of female authors and the total number of authors as well as quadratic terms for the number of female authors alone and interacted with the total number of authors (see Table S13 in the supplementary material for more details; refer to the above description of results for more details).

Significance levels: '$p < 0.1$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

team (e.g., adding one female author to an "all male" team may have a higher effect than adding one female author to a research team that already has five female authors). The coefficients for the quadratic terms and the interaction term between the number of female authors and the total number of authors are not reported in Tables 9 and 10. Instead, their effect was integrated into the reported coefficient for the number of female authors (see Tables S11 and S13 in the supplementary material for further details on the coefficient of these omitted terms).

To assess whether our results are consistent with the effect of an additional woman on interdisciplinarity being dependent on the number of women already included in a research team, Figures 6 and 7 illustrate the odds ratio for the number of women as a function of the starting number of female authors for publications with, respectively, seven authors (using the modeling results for 6–10 authors, Table 9) and 13 authors (using the modeling results for 11–20 authors, Table 10). The negative slopes on both plots suggest that an additional woman may indeed have a larger impact in teams mostly constituted by men, at least for these groups of publications.

It is also worth noting that the odds ratios of the interaction term between the number of female authors and the total number of authors was very slightly, but significantly, smaller than 1 for the models presented in Table 9 (6–10 authors) and Table 10 (11–20 authors) (note that the odds ratios for this term are only presented in Tables S11 and S13 of the supplementary material). This suggests that the effect of women on interdisciplinarity diminishes as more authors are included in the research teams. Furthermore, while the odds ratio for the number of authors was significantly higher than 1 for the group of papers with 3–5 authors (Table 8), it was not significantly different from 1 for the group of papers with 6–10 authors (Table 9) and significantly smaller than 1 for the group of papers with 11–20 authors. There may thus be a maximum team size beyond which interdisciplinarity even starts diminishing.

In the multivariate modeling, the neutral (or slightly positive) effect of self-citations on interdisciplinary depicted in Section 3.3 becomes slightly negative, yet statistically significant, in
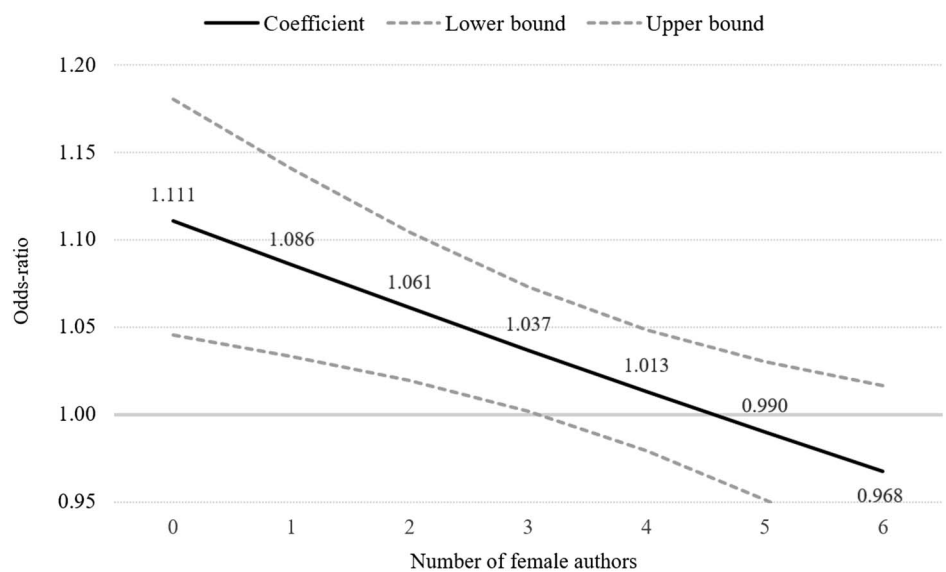
Figure 6. Effect of one extra female author on highly interdisciplinary papers as a function of the starting number of female authors. Refers to the estimated coefficients for papers with seven authors using MS2, RS, and GA # 1. The odds ratio for a starting number of two female authors is equivalent to that reported in Table 9. The lower and upper bounds refer to the 95% confidence interval.
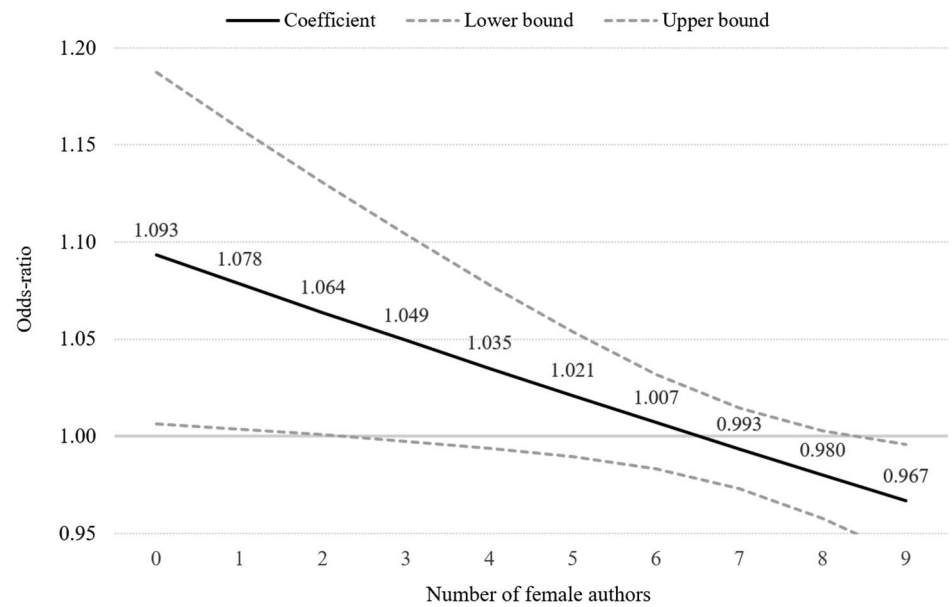
Figure 7. Effect of one extra female author on highly interdisciplinary papers as a function of the starting number of female authors. Refers to the estimated coefficients for papers with 13 authors using MS2, RS and GA # 1. The odds ratio for a starting number of four female authors is equivalent to that reported in Table 10. The lower and upper bounds refer to the 95% confidence interval.

nearly all groups of regressions (based on number of authors). This may be due to the inclusion of the length of a paper's reference list and its number of coauthors (two confounders of the effect of self-citations on interdisciplinarity) as control variables in the regression models. Using RS, each additional self-citation is associated with a statistically significant decrease in the range from −0.5% to −1.6% for groups of papers with 1–10 authors (i.e., groups 1–4 in Table 5). Only for papers with 11–20 authors is there no measurable effect. As was mostly the case for the effect of gender on interdisciplinarity (except for papers with 11–20 authors), the effect of self-citations is robust to changes in the model specification, the gender assignation rule, the diversity metrics for measuring interdisciplinarity, and the variable type used to express interdisciplinarity (see Tables 6–10; also see Tables S1–S10 in the supplementary material). In fact, the magnitude of the negative effect is even slightly stronger, ranging from −1.2% (odds ratio = 0.987) to −2.2% (odds ratio = 0.976), using DIV* to capture interdisciplinarity; in this case, the result is also statistically significant for the group of papers with 11–20 coauthors. It is also worth noting that the negative effect of self-citations on interdisciplinarity is systematically reduced, and sometimes even canceled, when interdisciplinarity is measured excluding self-citations (see Tables S1–S10 in the supplementary material).

This would partly explain why the effect size for gender (defined based on the presence and number of female authors on a publication) is systematically smaller under MS2 compared to MS1; in other words, part of the gender effect mediated through self-citations would be absorbed by the inclusion of self-citations in MS2. Per the summary presented in Table 5, and excluding papers with 11–20 authors, the effect size ranged from 6.9% to 24.8% in MS1 compared to 5.5% to 13.9% in MS2 (using the average number of authors and women in the group of papers with 6–10 authors). The largest effects under both model specifications were observed for the largest group of publications—that is, publications with 3–5 authors (50% of the study data set). As for self-citations, the magnitude of the effect is systematically stronger when using DIV* to capture interdisciplinarity instead of RS (Tables 6–10; also see

Tables S1 to S5 of the supplementary material). In that case, for papers with 3–5 coauthors, the magnitude of the relationship even increases for each additional female author (from 16.3% for one woman to 40.7% for five women).

Across all variants of the regression models (Tables 6–10; also see Tables S1–S10 of the supplementary material), less consistent results were observed for the variables related to seniority and prior publications. For single-author papers (Table 6), even though the odds ratios are significant for both controls, the magnitude of the corresponding effects are negligible. The same was generally observed for minimum seniority among a paper's coauthors and the volume of prior publications (regardless of how it is defined) (Tables 7–10). Interestingly, a negative and statistically significant effect was observed for maximum seniority within a paper's coauthors for all groups of regressions relying on publications with at least two authors (Tables 7–10); these effects were also consistent across all robustness tests (supplementary material).

4. DISCUSSION AND FUTURE DIRECTIONS

This study's goal was to assess the propensity of women, relative to men, to undertake interdisciplinary research in a large data set of ERA publications while accounting for a potential measurement bias mediated by self-citations, among other factors. If we consider that the number of prior publications is positively associated with self-citations, and that female researchers tend to accumulate fewer papers than men due to higher attrition rates and shorter careers (although this is gradually changing) (Mishra et al., 2018), any gender difference concerning the disciplinary diversity of a paper's cited references may be partly due to a measurement bias. This bias would be attributable to gender differences in self-citations mediated through gender differences in seniority and/or volume of prior publications. For example, self-citations may induce a downward bias in a paper's interdisciplinarity by reducing the balance of represented disciplines and the average distance between them, because the prior papers of an author are likely to be concentrated in one or a few closely related subfields.

In a context where funding for interdisciplinary research is gaining in importance to help solve the increasingly complex problems faced by modern societies (such as those being addressed through the UN SDGs), obtaining accurate measurements of interdisciplinarity by gender is highly relevant for funders in taking appropriate action(s) to level the playing field across genders. This is even more relevant given the conflicting evidence (Leahey et al., 2017; Rhoten & Pfirman, 2007; Science-Metrix, 2019) concerning the presence of women in interdisciplinary research, some of which could be biased by self-citations and other factors.

The key result of this study is that the presence of women in scientific publications is positively associated with interdisciplinarity even after controlling for a potential self-citation bias. As discussed in Section 2.5.3, testing for the presence of an association between gender and interdisciplinarity, while eliminating a potential bias mediated through self-citations, required advanced multivariate modeling to control for several confounders and mediators. Some of the mediators (self-citations, seniority, and prior publications) included in the full model specification (MS2) unfortunately have the potential to act as colliders that may reveal an association between gender and interdisciplinarity when there is in fact no such relationship. In this context, the simpler model specification (MS1), which only controlled for a confounder of the link between gender and interdisciplinarity (i.e., a paper's number of authors), served the purpose of assessing whether the relationship between gender and interdisciplinarity remained significant upon the exclusion of potential colliders (i.e., seniority, prior publications and self-citations) and the number of references (a confounder of the effect of self-citations on

interdisciplinarity assumed not to mediate an effect of gender on interdisciplinarity). In other words, MS1 estimated the total effect of gender on interdisciplinarity, including the portion mediated through other variables.

Based on the collective set of regression models estimated in this study, one can conclude that gender (defined as the presence and number of female authors) and interdisciplinarity are positively related on most models reported in this study. Apart from a few minor exceptions, odds ratios were systematically and significantly above 1 (i.e., positive regression coefficients) in the group of models involving up to 10 authors, including in our robustness tests (i.e., across model specifications, gender assignation rules, diversity metrics for measuring interdisciplinarity, and the variable type used to express interdisciplinarity [i.e., as a binary or continuous outcome variable]). For papers involving 11–20 coauthors, we could not conclude a positive (or negative) association due to inconsistent findings across our robustness tests. For example, the results were not significant when using RS, the main indicator of interdisciplinarity employed in this study, to identify highly interdisciplinary papers.

For groups of papers with 10 authors or fewer, the odds ratios for gender were systematically larger with MS1 (excluding potential colliders) than with MS2, where a portion of the total effect of gender on interdisciplinarity mediated through other factors was absorbed by the control variables. That said, the portion of the total effect mediated through other variables, including a potential measurement bias owing to women self-citing more than men, was not enough to fully remove a direct effect of gender on interdisciplinarity, which remains of an appreciable size (from 5.5% to 13.9% increase relative to a baseline of 10% using RS and from 10.2% to 40.7% using DIV*; effects measured using GA #1 and interdisciplinarity as a binary outcome; note that for papers with 6–10 authors, the measured effect is reported for the average team size (i.e., seven-author publications) and average number of women already figuring in the team (i.e., two).

Recall that there is a possibility of reverse causality between gender and interdisciplinarity induced by the inclusion of the confounder (i.e., number of authors). As the confounder cannot be omitted from either model specification, the relationship between gender and interdisciplinarity must be interpreted cautiously, emphasizing its strength over its causal direction. It is also useful to recall that the total and direct effects of gender on interdisciplinarity, as respectively captured with MS1 and MS2, voluntarily accounted for differences in gender participation, citation practices, and database coverage across subfields. These differences were absorbed by the subfield fixed-effects.

After controlling for the length of a paper's reference list and its number of coauthors—two confounders of the effect of self-citations on interdisciplinarity—the broad range of estimated regression models (using MS2) showed a statistically significant and negative effect of self-citations, except for papers with 11–20 authors when relying on RS to measure interdisciplinarity. Although this effect was small (RS: [–0.5%, –1.6%] excluding papers with 11–20 authors; DIV*: [−1.2%, −2.2%]), its consistency across all our robustness tests—combined with the fact that it was systematically reduced, and sometimes even canceled upon measuring interdisciplinarity without self-citations—suggests that self-citations may indeed induce a slight negative bias on interdisciplinary measurements. In the group of papers with the largest number of authors, an effect may not be present with RS and may be weaker with DIV*. This could be due to the self-citations in large teams being more likely to originate from a diverse set of disciplines, reflecting the disciplinary background of the contributing authors.

Part of the positive and systematic effect of the number of women on interdisciplinarity detected with MS1 (i.e., controlling for a key confounder and excluding potential colliders)

may thus be mediated by this measurement bias because women generally self-cite less than men and would thus be less impacted by a negative effect of self-citations on interdisciplinarity. Accordingly, the total effect of gender on interdisciplinarity, as reported above using MS1, may be slightly overestimated due to the self-citations potentially mediating a small negative bias in measuring interdisciplinarity.

The regression coefficients obtained for the total number of authors (odds ratios below 1) suggest that there may be a tipping point beyond which team size is negatively associated with interdisciplinarity. This could be the case if there was an optimal team size for running interdisciplinary projects; beyond a certain size, teams would not effectively collaborate across disciplines due to, for example, ineffective communication across scientific cultures (Kuhn, 2000). The odds ratios for the interaction between the total number of authors and the number of female authors were also smaller than one, suggesting that the association between the number of female authors and interdisciplinarity is conditional on team size itself, resulting in little or no effect of gender on interdisciplinarity for papers with large research teams. This could explain why the measured effect of gender on interdisciplinarity was less consistent in the group of papers with 11–20 authors. Finally, it was also found that for papers with more than six authors, the effect of an additional woman on interdisciplinarity diminishes as the presence of men gradually decreases in the team.

Note that MS2 also included seniority and the prior volume of a researcher's publication portfolio as potential mediators of an effect of gender on interdisciplinarity. As discussed in Section 2.5.3, there are several mechanisms through which these two control variables could mediate such an effect, the one owing to a measurement bias already being accounted for by the inclusion of self-citations as a control variable. Statistically significant findings were only found in a consistent manner for maximum seniority. In that case, a negative and statistically significant effect was observed for all groups of regressions (including robustness tests) except those based on single-author papers. These results may indicate that more experienced researchers, not necessarily more prolific authors, are slightly less attracted/motivated by the prospect offered by interdisciplinary collaboration when well established in a given field of research. This would appear consistent with survey results by Rhoten and Parker (2004), excluding the principal investigator group, for which their sample size was rather small. With women being less well represented among senior researchers (Directorate-General for Research and Innovation, 2019), this effect may also contribute to the total positive effect of women on interdisciplinarity. Still, as was the case with self-citations, the magnitude of this effect is small (using a 10% baseline for interdisciplinarity, the effect ranged from −1.5% (odds ratio = 0.9836) to −3.4% (0.9626) for each additional year of experience relying on RS and GA #1).

In summary, the combined evidence from MS1 and MS2 provide evidence of a positive link between women and interdisciplinary research in the large-scale data set under investigation. This contradicts some of the prior literature on the topic (Leahey et al., 2017) but confirms results from other studies making use of different approaches—for example, using survey data (Rhoten & Pfirman, 2007). These results add to the body of literature that may suggest research performing and funding organizations should not be too concerned by women being potentially discouraged from taking part in interdisciplinary work. Nevertheless, further research would be warranted to better understand the potential impact of observed gender differences in interdisciplinarity on, for example, career progression. For instance, prior work has suggested that current evaluation practices underlying tenure decision may not be properly accounting for the specificities of interdisciplinary work (Gewin, 2014; Rhoten & Parker, 2004). Also, further studies relying on alternative sources of

information to capture interdisciplinarity (e.g., semantic analysis of full texts) could be useful to further test the robustness of this study's results.

Further research would also be warranted to confirm the negative effect of self-citations on interdisciplinarity. Indeed, the regression coefficient for the effect of self-citations on interdisciplinarity was analyzed in the context of a regression model designed to account for the potential confounders and mediators of an effect of gender, rather than of self-citations, on interdisciplinarity. While we believe we accounted for the main confounders (i.e., the number of references and number of authors) of an effect of self-citations on interdisciplinarity, additional confounders might have been overlooked.

AUTHOR CONTRIBUTIONS

Henrique Pinheiro: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing—original draft, Writing—review & editing. Matt Durning: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing—original draft. David Campbell: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, Writing—original draft, Writing—review & editing.

COMPETING INTERESTS

The authors are employees of Elsevier.

DATA AVAILABILITY

We have added extensive supplementary material to the article as robustness tests of the study's findings. The complete bibliographic records used in this study cannot be released due to copyright restrictions.

REFERENCES

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*. https://doi.org/10.7554/eLife.45374, PubMed: 31305239

Beaudet, A., Campbell, D., Côté, G., Haustein, S., Lefebvre, C., & Roberge, G. (2014). *Bibliometric study in support of Norway's strategy for international research collaboration*. Oslo, Norway: Prepared for the Research Council of Norway. Retrieved from https://www.science-metrix.com/sites/default/files/science-metrix/publications/bibliometricstudyinternational.final_.pdf

Berge, L. (2018). *Efficient estimation of maximum likelihood models with multiple fixed-effects: The R package FENmlm*. CREA Discussion Papers.

Blackmore, P., & Kandiko, C. B. (2011). Interdisciplinarity within an academic career. *Research in Post-Compulsory Education*, *16*(1), 123–134. https://doi.org/10.1080/13596748.2011.549742

Campbell, D., Deschamps, P., Côté, G., Roberge, G., Lefebvre, C., & Archambault, É. (2015). Application of an "interdisciplinarity" metric at the paper level and its use in a comparative analysis of the most publishing ERA and non-ERA universities. In *20th*

International Conference on Science and Technology Indicators. Retrieved from https://science-metrix.com/sites/default/files/science-metrix/publications/campbell_et_al_sti2015_short_paper_final_web.pdf

Campbell, D., & Struck, B. (2019). Reliability of Scopus author identifiers (AUIDs) for research evaluation purposes at different scales. In *17th International Conference of the International Society for Scientometrics and Informetrics (ISSI), 2–5 September 2019, Proceedings Vol. II* (pp. 1276–1287). Retrieved from https://issi-society.org/publications/issi-conference-proceedings/proceedings-of-issi-2019/

Chawla, D. S. (2016). Men cite themselves more than women do. *Nature*, *535*(7611), 212. https://doi.org/10.1038/nature.2016.20176, PubMed: 27414239

Chen, S., Arsenault, C., & Larivière, V. (2015). Are top-cited papers more interdisciplinary? *Journal of Informetrics*, *9*(4), 1034–1046. https://doi.org/10.1016/j.joi.2015.09.003

Directorate-General for Research and Innovation. (2019). *She Figures 2018*. Retrieved from https://op.europa.eu/en/publication-detail/-/publication/9540ffa1-4478-11e9-a8ed-01aa75ed71a1

Freeman, R. B., & Huang, W. (2014). Collaboration: Strength in diversity. *Nature*, *513*(7518), 305–305. https://doi.org/10.1038/513305a, PubMed: 25230634

Gewin, V. (2014). Interdisciplinary research: Break out. *Nature, 511*, 371–373. https://doi.org/10.1038/nj7509-371a, PubMed: 25035881

Huang, J., Gates, A. J., Sinatra, R., & Barabási, A. L. (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(9), 4609–4616. https://doi.org/10.1073/pnas.1914221117, PubMed: 32071248

King, M. M., Bergstrom, C. T., Correll, S. J., Jacquet, J., & West, J. D. (2017). Men set their own cites high: Gender and self-citation across fields and over time. *Socius: Sociological Research for a Dynamic World*, *3*, 237802311773890. https://doi.org/10.1177/2378023117738903

Kuhn, T. (2000). Barriers to interdisciplinary research and training. In T. C. Pellmar & L. Eisenberg (Eds.), *Bridging disciplines in the brain, behavioral, and clinical sciences*. National Academies Press (US). Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK44876/

Leahey, E., Beckman, C. M., & Stanko, T. L. (2017). Prominent but less productive. *Administrative Science Quarterly*, *62*(1), 105–139. https://doi.org/10.1177/0001839216665364

Malcolm, B. (2021). *ggdag: Analyze and create elegant directed acyclic graphs.* R package version 0.2.4. Retrieved from https://cran.r-project.org/package=ggdag

Mishra, S., Fegley, B. D., Diesner, J., & Torvik, V. I. (2018). Self-citation is the hallmark of productive authors, of any gender. *PLOS ONE*, *13*(9), e0195773. https://doi.org/10.1371/journal.pone.0195773, PubMed: 30256792

Pinheiro, H., Vignola-Gagné, E., & Campbell, D. (2021). A large-scale validation of the relationship between cross-disciplinary research and its uptake in policy-related documents, using the novel Overton altmetrics database. *Quantitative Science Studies*, *2*(2), 616–642. https://doi.org/10.1162/qss_a_00137

Porter, A. L., & Rafols, I. (2009). Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics*, *81*(3), 719–745. https://doi.org/10.1007/s11192-008-2197-2

PPMI & Science-Metrix. (2018). *ERA Monitoring Handbook*. https://doi.org/10.2777/764148

Rhoten, D., & Parker, A. (2004). Risks and rewards of an interdisciplinary research path. *Science*, *306*(5704), 2046. https://doi.org/10.1126/science.1103628, PubMed: 15604393

Rhoten, D., & Pfirman, S. (2007). Women in interdisciplinary science: Exploring preferences and consequences. *Research Policy*, *36*(1), 56–75. https://doi.org/10.1016/j.respol.2006.08.001

Rivest, M., Vignola-Gagné, E., & Archambault, E. (2021). Article-level classification of scientific publications: a comparison of deep learning, direct citation and bibliographic coupling. *PLOS ONE*, *16*(5), e0251493. https://doi.org/10.1371/journal.pone.0251493, PubMed: 33974653

Science-Metrix. (2019). *Bibliometric study for the evaluation of the NSERC Discovery Research Program — Summary report*. Ottawa, Canada: Prepared for the Natural Sciences and Engineering Research Council of Canada.

Smith-Doerr, L., & Croissant, J. (2016). Gender equity and interdisciplinary collaboration. Retrieved February 25, 2020, from https://items.ssrc.org/interdisciplinarity/gender-equity-and-interdisciplinary-collaboration/

Zhang, L., & Leydesdorff, L. (2021). The scientometric measurement of interdisciplinarity and diversity in the research portfolios of chinese universities. *Journal of Data and Information Science*, *6*(4), 13–35. https://doi.org/10.2478/JDIS-2021-0027